Engineering plant-microbe communication for synthetic symbioses

by

Tyler Daniel Toth

B.S. Biomedical Engineering
University of Arizona (2016)

Submitted to the Department of Biological Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biological Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2022

Signature of Author ……
Department of Biological Engineering
August 16, 2022

Certified by……
Christopher A. Voigt
Professor of Biological Engineering
Thesis Supervisor

Certified by……
Angela M. Belcher
Department Head and Professor of Biological Engineering
Professor of Materials Science and Engineering
Thesis Supervisor

Accepted by……
Eric J. Alm
Professor of Biological Engineering
Professor of Civil and Environmental Engineering
Chair, Graduate Program Committee

Engineering plant-microbe communication for synthetic symbioses

by

Tyler Daniel Toth

Submitted to the Department of Biological Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biological Engineering

Abstract

Current agricultural practices will not be able to keep pace with a growing population in a changing climate. A better understanding and increased reliance on the symbiotic relationships between plants and their resident microbiome will be crucial in addressing this challenge. In addition, these microbes provide a promising location for genetic engineering efforts and should be viewed as an integral part of an engineerable agricultural system. This holistic view means that desired functions can be performed in the host most suitable for the task, reducing toxicity due to resource limitations, and potentially easing regulatory concerns. Attaining this view, however, requires synthetic forms of communication between plants and microbes that are orthogonal to native signaling pathways, able to diffuse through the complex soil environment, readily produced from common cellular precursors, and easily sensed at low concentrations.

In this thesis, I review the genetic parts and regulators available for use in plants and the types of natural and engineerable plant-microbe communication. I then highlight the potential use of plant-to-microbe communication to control gene expression in engineered soil bacteria – specifically to ensure that the energy-intensive expression of nitrogenase only occurs when microbes are near a plant. Finally, I share the creation of an engineered form of microbe-to-plant communication. We engineered plants with the ability to sense and respond to bacterial quorum signals separate from native responses. In addition, we show that the *p*-coumarate homoserine lactone (pC-HSL) sensors can respond to pC-HSL biosynthesized by *Pseudomonas putida* grown in proximity to the plants. To the best of our knowledge, this is the first demonstration of engineered microbe-to-plant communication using a small molecule. We were also able to place the biosynthesis of pC-HSL under the control of various sensors and use an engineered consortium to perform logical operations on multiple environmental inputs. This engineered form of synthetic symbiosis lays the foundation for using microbe-to-plant communication to perform tasks such as monitoring soil nutrients, sensing pathogens, or detecting environmental contaminants.

Thesis Supervisor: Christopher A. Voigt
Title: Professor of Biological Engineering

Thesis Supervisor: Angela M. Belcher
Title: Department Head and Professor of Biological Engineering

Contents

List of Figures

List of Tables

1 Overview of synthetic biology in agriculture

This chapter is composed of portions of a review article in progress covering synthetic biology in the agricultural sector cowritten by myself, Angelina Nou, Qiguo Yu, Yonatan Chemla, Philip Clauer, Kwan Yoon, and Christopher Voigt. The sections below were selected to highlight two themes: 1) the current state of synthetic biology in plants and 2) the types of natural and engineerable interkingdom communication.

1.1 Introduction

Traditional plant breeding and precision farming are not keeping pace with demands of global population growth[1, 2]. In addition to food, the growing bio-economy will place increasing pressure on agriculture for energy, chemicals, materials, and medicines that were previously derived from non-renewable feedstock[3]. Higher yields need to be obtained from less land, without destroying the environment. The United Nations projects that by 2050, crop yields must increase by 50% compared to 2013, but the reality is that productivity is stagnating across staple crops, including maize, rice, wheat and soy[4, 5]. Over the last 30 years, genetic engineering has improved yields while lowering carbon emissions and has reduced impact of uncertainties such as pests and weather. However, the widespread use of

engineered crops and microbes has been slowed by costly regulatory hurdles, international disparity in rules, and public acceptance. As a result, the ancient blights of agriculture remain; it is estimated that we face 50-70% reduction in yield due to abiotic stressors (drought, heat, salt, etc) [6], 17-30% due to pathogens and pests[7], and 14% due to production and supply chain loss and consumer waste[8]. As human-caused climate change continues, many of these stressors will be exacerbated forcing even more reliance on genetically engineered systems to potentially meet these challenges. As if the agricultural impacts were not enough, genetically engineered plants could also serve functions as sentinels, reporting information from the field, aid in bio-remediation of pollutants, and provide access to fresh water.

Synthetic biology encompasses a suite of modern tools for the design and genetic construction of living organisms. While developed with model species, such as *Escherichia coli* and *Saccharomyces cerevisiae*, the technologies are being ported to species of agricultural relevance, from plants and soil microbes to insects and viruses [9-11]. Crop genomes are more easily designed to carry multi-gene traits at precise locations in their chromosomes, thus simplifying trait stacking. Synthetic regulatory networks reduce the load of traits on cells by turning them on only when needed; for instance, responding to a pest attack by releasing a pulse of insecticide. Plants grow slowly, making it difficult to build and evaluate complex designs. This could be aided by advances in computer aided design to improve the likelihood of success and automated rapid prototyping systems that could evaluate thousands of designs in parallel using cell-free or protoplast systems. Machine learning, combined with –omics and rapid prototyping, may enable the mapping of these results to performance in a crop plant in the field. This review looks at the emerging application of these concepts to crop plants and soil bacteria.

These advances extrapolate to a future where agricultural solutions involve interlocked species, involving engineered plants, microbes, and insects, all working in concert to maximize yields and fend off threats. One can imagine planting a seed that is coated with a myriad of engineered bacteria and fungi that communicate with each other and the seedling as it germinates and grows. Functions can be distributed across this integrated agricultural system depending on where they are best implemented. For instance, a plant-based sensor may respond to a pathogen by signaling to a bacterium in the root to make and deliver an antibiotic that is more easily built by prokaryotic chemical synthesis machinery. This cell-cell communication could be integrated with engineered insects, called to use their mobility to bring functions to an otherwise immobile system. The plants and microbes could be programmed to communicate directly with electronic sensors, for example mounted on unmanned aerial vehicles, to report information about the state of the crops. Information could also be fed to the crops, perhaps tweaking genes to resist an upcoming weather event.

This review covers the nascent application of principles of synthetic biology to agriculture and extrapolates to the near and far future. The first section describes the impact of genome editing, early

collections of genetic parts and the construction of the first genetic sensors and circuits towards the creation of engineered plants that can respond to their environment. Then, the review covers the engineering of other species of relevance, from soil bacteria to phage and insects and how these could be connected into integrated systems. To meet the demands of the future, agricultural systems will be highly-engineerable, where we have full understanding and design control over not just the genome of the crop, but all the bacteria and fungi and supporting insects.

1.2 Plant design and construction

New tools to manipulate genomes have radically changed the scale and ambition of plant genetic engineering projects. DNA synthesis and assembly can be used to realize a design, encompassing megabases in extreme examples, with every base pair specified. Genome editing can target these designs to specific locations in the genome, as well as knockout or mutate native sequences. Within a plant, it is becoming viable to divide engineering efforts across nuclear chromosomes, the plastid, and mitochondrial DNA, selecting the location best suited for the desired trait. Beyond traditional traits, engineered functions can redirect carbon flux through central metabolism, build synthetic regulatory networks to control the plant's response, or introduce multi-enzyme pathways from other organisms to build flavors, pharmaceuticals, materials, or other products. While slow to be developed for plants, libraries of genetic parts are becoming available that allow a designer to control protein expression and localization. Artificial regulatory networks comprised of sensors and circuits can control when traits are turned on to actuate a dynamic response, such as during development, or only turn on a trait when needed, thus reducing resource commitment. Genetic engineering projects are getting increasingly complex – combining regulatory control, metabolic flux analysis across organelles, and controlling plant development – and this will ultimately necessitate advanced and integrated CAD software to balance constraints. Further, plants are notoriously slow growers, and accelerating the design cycle will require prototyping methods to experimentally validate designs in high-throughput before committing to the construction of a plant.

1.2.1 Engineering plant genomes

Precision genome editing has been facilitated by reprogrammable DNA binding domains that target a DNA-cutting nuclease to a specific location in the genome. Many variants of this approach can be used to knockout genes, make specific base pair changes, or drop payloads of synthetic DNA into the genome. These components can be delivered by plant viruses or *Agrobacterium*, either as a laboratory technique or edit plants in the field.

1.2.2 Editing chromosomes

Editing tools delete, modify, or insert DNA at a precise location in the genome[9, 12]. DNA recognition and cleavage can be performed by meganucleases, zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), or CRISPR/Cas systems, all of which have been shown to work in plants[13-16]. Common to each of these proteins are domains that can be reprogrammed to theoretically target any desired DNA sequence. In practice, however, they all have limitations. Meganucleases and ZFNs are particularly difficult to design, with only a handful of examples of their use in crop plants[13, 17-22]. TALENs are more easily designed by combining domains, each of which contributes modularly to the DNA sequence that is recognized[23, 24]. CRISPR/Cas9, however, has rapidly become the method of choice for engineering plants due to the simplicity of designing a sgRNA to target almost any DNA sequence[16, 25, 26]. In a recent survey of genome edited crops, 161 traits have been introduced using Cas9 and 20 using TALENs[27].

Despite its widespread use, Cas9 is still partially limited by its large size, off-target effects, and the requirement of a PAM sequence (protospacer adjacent motif, short DNA sequence adjacent to targeted DNA region for cleavage by CRISPR/Cas system) at the target site. There has been rapid development in the CRISPR field to find or engineer alternative machinery that addresses these limitations. Once example is CasΦ, a minimal CRISPR/Cas system identified from a bacteriophage that is half the molecular weight of Cas9, has more flexibility in the sites it can integrate, and has been shown to function in plant cell protoplasts[28]. Variations of CRISPR systems can naturally have alternative PAM sequences but can also be evolved to further expand the targeting capabilities; thus, the number of sequences that can be targeted is increasing. For example, CRISPR/Cas12a has different PAM specificity than Cas9 and was applied to edit the promoter of the *CsLOB1* gene in Duncan grapefruit to increase its resistance to citrus canker, a feat not possible with Cas9[29].

There are many examples of crops being improved with simple gene knockouts. Prior to editing technology, screening campaigns identified many such examples, but the techniques required transgenes or left foreign DNA in the genome. These crops are treated by the USDA as GMOs, which increases development cost and time. For example, the Flavr Savr tomato exhibits extended storage life via suppression of the polygalacturonase (*PG*) gene by RNAi [30]. The same modification can be recapitulated by CRISPR/Cas to generate a tomato *PG* mutant[31]. Crucially, genome editing techniques that leave no foreign DNA behind drastically simplify the regulatory process. In 2016, the first CRISPR-edited food product to gain this status was an anti-browning white mushroom made by knocking out polyphenol oxidase (*PPO*)[32]. The first genome-edited crop to enter the US food supply was a high-oleic soy with improved oil stability, developed by knocking out two fatty acid desaturase 2 genes (*FAD2-1A* and *FAD2-1B*) with TALENs[33]. This crop entered the market in 2019 after the USDA decided not to regulate it and other crops that do not contain foreign DNA[34, 35]. This was followed by a waxy corn, high-oil-content camelina, and

drought tolerant soybean[35]. More recently, yield-related traits can be enhanced by generating different *cis*-regulatory alleles using CRISPR/Cas in maize and tomato[36-38].

Genome editing tools can also make single nucleotide replacements in the genome. Oligonucleotide-directed mutagenesis, an early technique used to make herbicide resistant crops, relied on double stranded DNA:RNA oligos containing the desired nucleotide change within the target sequence, but suffered from low conversion efficiency[39]. "Base-editors" make single targeted replacements by use of a nuclease-deactivated Cas9 (dCas9) to position an enzyme capable of a C-to-T (cytosine base editor; CBE) or A-to-G (adenine base editor; ABE) mutation[40]. Herbicide resistant rice, wheat, maize, cassava, and watermelon have been created by making mutants of the enzymes targeted by the herbicides[41-46]. Base-editing has also been used to substitute a single base pair in the *ALC* gene to create a nonsynonymous amino acid change which increased the shelf life of tomato[47]. The recently developed "prime editors" can generate all 12 types of base conversions without requiring double-strand breaks or donor DNA templates. It relies on the prime editor itself, a fusion protein between Cas9 nickase and reverse transcriptase-RT and the prime-editing guide RNA (pegRNA), containing the primer binding site (PBS) and the RT template. The desired edits in the RT template can be reverse transcribed and inserted into the target site. Herbicide resistance rice and maize have also bene generated using prime editing, though editing efficiency remains low.

Multiplexed genome editing in plants allows simultaneously editing multiple targets. As an example, 107 of 109 members of the caffeic acid *O*-methyltrasnferase (*COMT*) gene family were mutated with a single pair of TALENs in sugarcane to improve the bioethanol production from lignocellulosic biomass. CRISPR/Cas system offers more flexibility since gRNAs can be expressed as a polycistronic transcript and processed into mature gRNAs using tRNAs, ribozymes or endogenous nucleases. This approach has been applied to improve lycopene content in tomato fruits using six sgRNAs, decrease the glycoalkaloids in potato tuber using nine sgRNAs, or eliminate the glycan in tobacco by mutating 14 loci[48-50].

1.2.3 Chromosomal insertion of synthetic DNA

Payloads of synthetic DNA can be inserted into the genome, for example to insert a trait involving a recombinant enzyme or pathway. Previously, methods to do this inserted the payload at multiple random locations, and plants would be screened for those with the desired phenotype and no discernable defects[51-54]. This complicates stacking because genes must be added serially and screened afresh leading to different insertion locations than the plants developed with individual traits. Because of this, to our knowledge, the most traits put into a single crop was seven in maize (three herbicide tolerances, two insect resistances, one antibiotic resistance, and one sugar metabolism)[55]. In contrast, by using gene editing nucleases, payloads

can be inserted at a specific location using nonhomologous end joining (NHEJ) and homologous recombination (HR) at a nuclease cut site. These insertions can be used to change the regulatory control of a gene; for example, maize drought tolerance was improved by replacing a native promoter with a stronger one to increase the expression of a negative regulator of ethylene responses[56].

A challenge with making large insertions is obtaining sufficient concentrations of template DNA inside the plant cell to trigger NHEJ or HR. One approach to address this challenge uses the geminivirus replicon to amplify the template DNA within a cell, which increases HR efficiency in potato, tomato, wheat, and rice[57-60]. Alternatively, chemically modified double-stranded oligodeoxynucleotide (dsODN) can be used to facilitate NHEJ. These dsODNs have been used to insert a 60 bp translation enhancer and a 2 kb promoter in rice as proof-of-principles and offers the potential to insert complete protein-coding genes[45, 61-63]. The efficiency of large insertions can be further increased by encoding a transposase/integrase into the sgRNA, but to date this approach has only been demonstrated in microbes[64-66].

The insertion of a large DNA payload into the genome can disrupt native genes and, conversely, the payload will be sensitive to the local chromosomal environment and background transcription. The insertion of "landing pads" into the chromosome first, into which the payload is later added, is one way to increase reliability. Landing pads in microbes, yeast, and mammalian cells has been created using two steps[67-69]. First, random insertions are screened to determine a chromosomal location from which high expression is achieved. These sites are evaluated for their impact on the host, for example by measuring changes in growth or more thoroughly by using RNA-seq to ensure that host gene expression is not impacted. Second, the landing pad is insulated by flanking it with terminators to block transcription into or out of the landing pad, factors to open up the chromatin, and insulators to protect against nucleosome changes[68]. Sometimes, insertion sites are included that are specific to the payload delivery method (phage integrase or transposase site). Creating landing pads in plants would allow for more predictable expression levels when a gene is inserted at that position. New classes of insulators are needed to decouple expression from neighboring enhancer sequences and block expansion of DNA silencing from neighboring regions[70].

Another way to isolate transgenes is to move them onto an artificial chromosome. In other eukaryotes, stable artificial chromosomes can be made by reintroducing centromere DNA to DNA containing genes or pathways of interest[71-74]. This does not work in plants because their centromeres contain large arrays of repeats, sometimes as large as a megabase[75] and putting these repeats on an artificial chromosome does not necessarily generate the kinetochore required for spindle fibers attachment during cell division. Therefore, a different approach has been taken to build plant artificial chromosomes based on the telomere-mediated truncation strategy originally developed in yeast[76]. For example, mini-chromosomes have been obtained from truncated A and B chromosomes in maize by the introduction of a transgene cassette harboring *Arabidopsis* telomeric sequence[77-79]. The artificial chromosome can have a

recombination site (e.g., *loxP*) to act like a landing pad to facilitate the insertion of transgenes[76]. To date, plants containing stable artificial chromosomes include rice, barley, wheat, *Arabidopsis*, and *Brassica*[80-85].

There are technical limitations on the size of a DNA payload that can be delivered to a plant cell. There have been a few successful examples of the delivery of large DNA sequences, for example, anthocyanin- and astaxanthin-rich rice were created by transferring ten (31 kb) and four (15 kb) biosynthetic genes from different organisms[86, 87]. Wheat resistant to the fungal pathogen *Puccinia graminis* f. sp. *Tritici* (*Pgt*) was generated by introducing a 37 kb T-DNA construct encoding five resistance genes into a single locus[88]. Two pathogen resistance gene clusters from *Arabidopsis* were fused using homologous recombination in yeast (20 kb) and transformed into soybean, demonstrating the possibility of transferring entire chromosomal regions between species[89]. While the largest DNA fragment delivered into a plant nucleus via *Agrobacterium*-mediated transformation is relatively large at 150 kb (a piece of human genomic DNA)[90], development of new vectors with increased carrying capacity will be essential.

1.2.4 In situ transformation

There are scenarios where it could be advantageous to deliver genetic material to a crop after it has been planted, for example to deliver resistance to an emerging insect pest. One way to do this would be to hijack plant viruses to deliver the payload in the field.

Autonomously replicating viral vectors based on infectious plant viruses have been used to express foreign genes in plants [91]. Plant viral vectors confer high expression yield of desired product and enable the rapid screening of multiple construct variants in different host species or genotypes and particular plant growth stages or even specific plant organs [92]. Engineered plant viral vectors to promote the transition of plants into desired developmental stages, like early flowering to accelerate breeding programs, was achieved by expressing FLOWERING LOCUS T (FT) [93]. Plant viral vectors have also been used to study and control pests by inducing RNAi of essential genes in the pest upon feeding [94-96] or lading pesticides onto virus nanoparticles[92]. Direct expression of antimicrobials or fusion of these peptides onto the viral capsid protein can be used to rapidly evaluate antimicrobial activity and increase plant innate immunity against pathogens[97-100]. Plant viral vectors have also been used to express sgRNAs for Cas9-based editing[101, 102].

Biocontainment of engineered recombinant viruses is essential since plant viruses can cause considerable economic losses in crops and infectious recombinant viruses can pose a risk for escape into the environment. Deconstructed viruses lacking one or more essential genes which are complemented only in a transgenic host provide a means of preventing a recombinant virus from infecting other hosts[103, 104]. Alternatively, mutating or deleting essential genes for virus transmission has been shown to be an effective biocontainment strategy[105].

1.2.5 Editing the plastid and mitochondrial genome

Mitochondria and plastids are prokaryotic in origin and share similar genome organization and expression machinery with bacteria. As such, they have potential for expressing bacterial pathways that would be incompatible with nuclear expression due to differences in genetic organization and regulation[106]. They also have enzyme cofactors, such as metal clusters, that are not present elsewhere in the cell, contain higher concentrations of some core metabolites, and have high (chloroplast) or low (mitochondria) oxygen tensions, which are desirable for some enzymes[107].

Plant leaves can have up to 100 chloroplasts per cell and 50 copies of the chloroplast genome (plastome) per chloroplast[108]. The plastid genome is relatively small, encoding about 100 genes, but can range from 19 to 217 kb across species, implying that it can carry more genes without overloading the cell[109]. When a transgene is introduced into the plastid, very high levels of expression can be achieved (>70% of the plant's total soluble protein)[110]. Maternal inheritance of the plastome mitigates environmental dispersion of transgenes through pollen. The chloroplast also lacks RNAi machinery, making it possible to express dsRNA, which has been used to confer insect resistance[111]. Non-green plastids, such as carotenoid-rich chromoplasts and starch-accumulating amyloplasts, do not host photosynthesis, are less sensitive to foreign product accumulation, and thus provide a promising location for metabolic engineering[112].

Despite many advantages, the difficulty of engineering the plastome currently limits its application in agriculture. Plastid engineering has not been demonstrated in any cereals and has only been successful in a limited number of agriculturally relevant crops: tobacco, potato, tomato, lettuce, soybean, rapeseed, carrot, cabbage, sugar cane, sugar beet, eggplant and cauliflower[113-121]. The large number of chloroplasts and plastids per cell makes it difficult and time-consuming to reach homoplasmy, in which all plastids contain the transgene[122]. Transplastomic rice have been obtained but were not in the homoplasmic state[123, 124]. Major cereal crops are recalcitrant to chloroplast transformation primarily due to the lack of sensitivity to the selection agent spectinomycin, which is an inhibitor of plastid translation. Another problem is that recombinant gene expression commonly has a negative impact on plants by interfering with photosynthesis. Non-green plastids, such as chromoplasts in tomato fruits and amyloplasts in potato tubers, provide a less-explored location for carrying a recombinant function as expression in these organelles does not impact photosynthesis.

Current chloroplast genome editing relies on the homologous recombination events to introduce foreign genes of interest along with a selectable marker gene into the endogenous plastome. Recently, the cytidine deaminase domain of *Burkholderia cenocepacia* DddA toxin was fused with the DNA-binding domain of a TALE array to to carry out the C to T conversion in single chloroplast gene (*psaA, psbA, 16S rRNA, rpoC1, atp6,*) in rice, Arabidopsis, lettuce, and rapeseed [125-127]. Alternatively, the transgene can be

delivered via an episomally replicating plasmid that leaves the native plastid genome untouched. The plasmids are amplified within the chloroplast through the replication origin derived from dinoflagellate plastid genome [128], tobacco native extrachromosomal element (NICE1) [129, 130], or with the help of a DNA replicase that recognize and replicate specific sequences flanking the transgene. However, the maintenance of the self-replicating plasmids when growing without selection pressure and its inheritance into subsequent generations via seeds remains challenging.

Mitochondria, the powerhouse of the cell, also contain independent genomes that evolved from prokaryotes. While mitochondria have been engineered in yeast, eukaryotic algae, and mammalian cells[131-136], the mitochondria from plants have only been transformed after isolating them from the plant cell[137, 138]. The small size of the organelle and lack of effective selectable marker genes are major obstacles for plant mitochondria transformation. However, instead of directly transforming mitochondria, gene editing nucleases can be expressed from the nuclear genome and directed into the mitochondria with signaling peptides [139, 140]. Nucleus-encoded TALENs fused to these mitochondrial signaling peptides have been used to knock out mitochondrial genes in rice and *Brassica* to cure cytoplasmic male sterility[140]. With the same approach, *atp6-1* and *atp6-2*, encoding ATP synthase subunit 6 was also successfully deleted in *Arabidopsis* mitochondria[141]. In addition, the DddA cytosine base editor has been used to edit *atp6* and *rps14* gene in lettuce and rapeseed, respectively[127, 136, 142, 143]. Collectively, the chloroplast and mitochondria provide additional locations with potentially desirable features for future engineering efforts .

1.3 Plant genetic system design

Abstraction is a core principle of synthetic biology intended to aid complex design projects[144]. Genetic parts are the most basic units, representing DNA sequences of a minimal fundamental molecular biology function, like a promoter. Large libraries of reliable genetic parts aid in precisely tuning expression levels, for example to balance flux through a metabolic pathway or facilitate the construction of large designs. Genetic devices are assemblies of parts that have a more abstract but functionally meaningful purpose, such as a sensor or metabolic pathway. Devices are assembled into the system, which represents the complete recombinant program introduced to the cell. Previous plant genetic engineering projects have not necessitated this structure as they were relatively simple, but as they get more complex over time, such hierarchies will be valuable in organizing designs. This section describes current genetic parts used in plants and some of the complex plant genetic systems that they have been used to build.

1.3.1 Plant genetic parts

Defining a genetic part for a plant can be more complex than other organisms. A part should be discrete, meaning it has a clear nucleotide beginning and end, and have a single function, irrespective of genetic context. These are often not the case for natural regulatory parts in plants; for example, promoters can have enhancers many kilobases upstream, functions can overlap, and chromosomal effects and silencing are often position dependent. Most of plant genetics will not conform to this idealized abstract structure, but that is okay – the point of the field is to simplify and collate synthetic genetics to facilitate engineering. Still, it is valuable to engineer part libraries that define parts, minimize context effects, and maintain them in public databases to facilitate widespread use. A notable example of this effort is OpenPlant [145].

Interestingly, one of the reasons that part libraries have been slow to develop for plants is due to the historical limitations of genome modifications. Because of the random nature of genome insertions, a single gene with one promoter and terminator could be transformed and it would insert into many chromosomal locations. These different locations lead to different levels of expression and the after-the-fact selection of a plant that has the desired phenotype is effectively a directed evolution experiment[146, 147]. However, as genome editing facilitates single, defined integrations it is now important to be able to vary expression levels through part selection. Libraries of genetic parts can be obtained using bioinformatics to search for lists of natural sequences ("part mining"), computational design, or randomizing a scaffold[148-151]. Part characterization is also more difficult in plants because of the difficulty in performing high-throughput experiments, variability in expression due to cell and tissue type, and fluorescent interference of chlorophyll with common expression reporters [152, 153].

Currently, the relatively high cost and long timespan required to genetically engineer plants make it difficult to say, build hundreds of lines just to characterize a small library of constitutive promoters. Large part libraries have been developed in other organisms using multiplexed DNA construction and high-throughput characterization strategies, such as flow-seq, where a cell sorter is used to isolate parts by their activity, followed by deep sequencing to determine the parts. As an example, this has been used to characterize over 12,000 *E. coli* promoters and RBSs in a single experiment[148, 154]. While this approach would obviously be unsuitable for screening whole plants, protoplasts, or individual plant cells with enzymatically removed cell walls, provide a promising way to scale throughput and have been used to characterize over 100 synthetic promoters in *Arabidopsis* and sorghum[155]. Protoplasts can also be sorted using fluorescence-activated cell sorting, although to our knowledge, this has not yet been applied to *in planta* part discovery[156, 157].

1.3.2 Transcriptional parts

In plants, transcription is regulated by numerous proteins which bind to specific DNA sequences to alter chromatin states, interact with general transcription factors, and recruit RNA polymerase II for

transcriptional initiation[158]. The minimal, also known as core, promoter is the DNA sequence that lies directly upstream of the transcription start site and acts as the binding site for the transcription pre-initiation complex, including RNA polymerase II[159]. The core promoter typically contains the TATA-box, initiation and/or downstream promoter elements, and is regulated with proximal and distal regions containing *cis*-regulatory motifs that act as binding sites for *trans*-acting transcription factor proteins[160]. For the purposes of abstraction, the strength of a promoter has been defined as the RNAP flux that emanates from it, although the expression of a protein reporter is often used as a surrogate. This is more complicated in eukaryotes because of all the steps after transcription that can alter expression levels, including capping, export, splicing, and polyadenylation of the 3' end[161]. A further complicating factor is that the strength of a promoter in a plant chromosome can be impacted by complex chromosome structures, nucleosome occupancy, pervasive transcription, and interactions with the nuclear pore complex[68].

Most engineering in plants to date has relied heavily on the use of a few well-established constitutive promoters to achieve high levels of expression, such as the 343 bp Cauliflower Mosaic Virus 35S (referred to as CaMV35S or 35S). Due to the fact that the 35S provides continuously high expression level of downstream genes and has a very wide host range, it has become the standard in which other plant promoters are typically compared[162]. The 46bp CaMV35S minimal promoter and its variations can be found in over 60% of all transgenic crops grown worldwide[162]. Other commonly used promoters are typically derived from plants or viruses that infect plants and can be tissue-specific and/or stress-inducible[163].

Repeated usage of a small set of promoters and terminators is not desirable due to potential problems with transfer-DNA (T-DNA) stability, integrity, and gene silencing[164]. Synthetic promoters may provide the solution to these issues. By screening different upstream *cis*-element binding sites and downstream minimal plant promoters, libraries consisting of hundreds of unique constitutive synthetic promoters with varying strengths have been created in *Arabidopsis*[165]. This approach was expanded upon by increasing the number of functional elements shuffled and creating a computational model to predict the strength of over 1,000 synthetic promoters with good correlation, a few of which resulted in expression comparable to CaMV35S although most of the designs had very low predicted strength ($R^2 = 0.7$ actual versus predicted)[166]. In addition, a subset of these promoters followed similar strength trends when tested in the *Brassica rapa* and *Nicotiana benthamiana* [166].

Many promoters do not provide the same expression level between species. To date, there have been few large-scale characterizations of promoters in different species. After testing 46 promoters in *Nicotiana benthamiana* leaves, *Medicago truncatula* roots, *Lotus japonicus*, and *Hordeum vulgare*, a core set of only seven was identified that showed functionality across the species and tissues tested[167].

New promoters, either constitutive, tissue-specific, or inducible, from different plants can be identified using microarrays or RNA-seq[168, 169]. The PlantProm database contains 576 plant promoters with

experimentally verified transcriptional start sites across 86 species including rice, maize, soybean, grape, *Medicago*, and *Arabidopsis*[170]. The largest collection of experimentally identified plant promoters is Plant Proteome DataBase (ppdb) with promoter positioning for ~27,000 (as of 2021) genes for *Arabidopsis*, as well as rice, *Physcomitrella patens*, and poplar[171]. Computational tools, such as TSSPlant, can be used to predict novel plant promoters using data from these large datasets[172].

Terminators stop RNAP progression and lead to the release of the transcript. In bacteria, their strength is quantified as the fraction of RNAP that are blocked[173, 174]. Different terminators can also lead to different 3'-hairpins, which impacts transcript stability[175]. In eukaryotes, their role can be more complex, and their strength can be promoter-dependent, possibly due to the formation of gene loops[176, 177]. Gene looping, or the direct interaction of the terminator with the promoter mediated by nucleic acid binding proteins, has been shown to influence expression. While the use of certain terminators commonly increases gene expression, typically through mRNA stabilization, the degree to which they increase expression varies by which promoter it is paired with which makes the standardization of parts more difficult[178]. Gene looping is believed to explain this variation through transcriptional memory[179], intron-mediated modulation of transcription[180], transcription directionality[181], reinitiation of transcription[182], and transcription termination[183].

In plants, the dominant feature affecting transcript stability is the terminator region[184]. For optimal transgene expression, constructs should have strong terminators to avoid the production of aberrant transcripts which lead to transcripts without polyadenylation, siRNA production, and ultimately transgene silencing[184]. Like promoters, terminator performance often depends on the host; therefore, viral- or pathogen-derived terminators, such as the NOS terminator from *Agrobacterium tumefaciens*, are commonly used due to their broad host range. Certain terminators have been found to be more effective in stabilizing transcripts, such as the HSP terminator which provides 2.5-fold increase in mRNA levels compared to the NOS terminator[185]. However, there remains a large degree of host specificity. Ten terminators screened in *Medicago truncatula* and barley had no correlation in terminator strengths between the two species highlighting the importance of identifying appropriate terminators for optimal transgene expression [167].

1.3.3 Translational parts

After a transcript is exported from the nucleus, a set of intrinsic features will dictate the efficiency of translation. In general, the upstream portion of mRNA contains the 7-methyl guanosine (m^7GpppN) cap which is typically required for translation. In additional, the 5' untranslated region (UTR) upstream of the translational start codon and 3' UTR downstream of the stop codon both play an important role regulating translation efficiency[186]. Finally, the poly(A) tail at the 3' end of the transcript recruits the protein complex that unfolds mRNA secondary structure and facilitates ribosomal assembly[187]. Secondary structures such

as hairpins can also block the progression of the ribosome thus reducing translation. Binding sites for RNA-binding proteins (RBPs) and small RNAs can also modulate translation. Unlike in prokaryotes, there is no ribosome binding site (RBS) to control the expression level. Still, modulation of these features and parts listed below are available to tune or enhance translation in plants.

Native plants genes are typically expressed as monocistrons and are not organized as operons. That being said, many eukaryotic mRNAs contain upstream open reading frames (uORFs) in the 5'UTR which typically reduce the translation the primary open reading frame (ORF)[186]. Incorporation of uORFs can be used to modulate the translation of transgene transcripts[188]. Ribosomal subunits can be recruited to an internal site in the transcript using an internal ribosome entry site (IRES) which bypasses the need for scanning and cap-dependent translation and enables polycistronic expression *in planta*[186]. Many plant RNA viruses rely on IRESs to recruit ribosomes without the standard protein translation initiation machinery[189]. The IRES from *Tobacco mosaic virus* promotes the translation of a second cistron to ~30% the levels of the first cistron in tobacco[190]. Several other IRES from plant viruses have been shown to function in planta and may provide additional non-canonical translational control of desired gene expression[187].

Translational enhancers can also be derived from many plant RNA viruses and are useful for attaining high levels of protein. While the efficacy of viral translational enhancers appears to be dependent on virus host range, commonly used enhancers, such as the tobacco mosaic virus (TMV) Ω enhancer, have been shown to recruit translational machinery similar to properly capped and polyadenylated transcripts[191, 192]. A 28 nt synthetic 5'-UTR was used to enhance gene expression in tobacco and cotton 10- to 50-fold depending on the tissue and was found to be more efficient than Ω in several tissues[193].

Inclusion of certain introns have been found to improve transgene expression; an effect known as intron-mediated enhancement (IME). While IME has been known to occur in plants for several decades, the exact mechanism is still unknown with hypotheses including increased rates of transcription, pre-mRNA processing, mRNA export, mRNA stability, and translation[194, 195]. Despite this, some design constraints have become clear - the introns must be placed in the correct orientation, within the transcribed sequence, and close to the transcription initiation start[196]. A few dozen introns have been used to modulate expression or restrict it to specific tissues[197]. However, due to the lack of mechanistic understanding, the use of IMEs to tune transgene expression is cumbersome – the same IME in different plants will often result in different expression patterns and thus more insight is likely needed before their broad use[196].

1.3.4 Post-translational parts

Post-translational modifications, such as peptide tags, provide an additional tool to tune the dynamics, localization, stoichiometry, and function of proteins of interest. For example, a temperature-responsive degradation domain from yeast was modified with two mutations to function in a temperature

range more permissive of plant growth resulting in the low temperature (lt) degron. When fused to the N-terminus of proteins, the lt degron results in temperature-responsive degradation of a protein-of-interest and was used to modulate native transcription factors involved in trichrome development and flowering time *Arabidopsis*[198, 199]. Certain degradation domains can also be protected by or induced by the presence of a small molecule, yet these systems have not been demonstrated in planta[200, 201].

Peptide tags have been used to localize proteins to the nucleus, plastids, and mitochondria[202]. Proteins are targeted to plastids by N-terminal transit peptides which commonly contain hydroxylated residues, such as serine, and generally lack acidic residues, resulting in a net positive charge [203]. The mitochondrial targeting peptide derived from CoxIV gene from *S. cerevisiae* as well as *Arabidopsis thaliana* F1-ATPase pFAγ subunit mitochondrial targeting peptide works in *Nicotiana benthamiana* transient leaf assay and *Arabidopsis* protoplasts[204, 205]. Nuclear import of proteins is essential for native gene regulation and is often highjacked by pathogenetic viruses and bacteria. Proteins of interest can be localized to the nucleus using nuclear localization signals (NLSs) from these sources, such as the Simian Virus 40, which are recognized by the native nuclear transport proteins.

Polycistronic cassettes can be imitated by placing cleavable peptide sequences between genes, thus providing an additional tool to control the ratios between desired genes. Incorporation of the self-cleaving 2A peptide from foot-and-mouth disease virus between two genes of interest allows constant expression levels, though it leaves an 18aa tail on the upstream protein which may influence protein function[206]. However, a synthetic 2A sequence was used to express two carotenoid biosynthetic genes under a single rice globulin promoter resulting in New Golden Rice that is less susceptible to transgene silencing (due to needing only a single promoter and terminator to express two genes)[207]. Alternatively, site-specific proteases, commonly from potyviruses, can be used to release individual components from a large polyprotein. This approach was used to reduce the number of *Klebsiella oxytoca* nitrogenase genes required for activity in *E. coli* from 18 down to 5 large polyproteins with 8 Tobacco Etch Potyvirus (TEV) sites[208]. Potyviruses are the largest group of plant infecting RNA viruses and thus there are an abundance of potential proteases that function in plants though they have not yet been widely adopted as a plant genetic part[209, 210].

Targeted changes to the epigenome can be accomplished using fusions of DNA targeting enzymes with methylases or demethylases. Synthetic zinc fingers have been used for targeted demethylation[211] or to target methylation to the binding element of a disease-causing TAL to provide resistance for the plant[212]. Using the dCas9-SunTag system can allow for even easier targeted demethylation to modulate native or transgene expression[213].

Gene silencing mediated by small RNAs (sRNAs) play an important role in regulating gene expression in plants, with transgenes often being more susceptible to silencing machinery[184]. Transgenes generated only using the coding sequence of a gene (without the presence of introns) and under non-plant

regulatory regions, such as viral promoters and pathogen-derived terminators, resemble viral or bacterial genes and can lead to siRNA production and ultimately transgene silencing[214]. Transgenes designed with plant-derived 5' and 3' regulatory regions, plus introns of the tobacco RuBisCO, were shown to be less prone to post-translational and transcriptional gene silencing[131-136]. Alternatively, simultaneous expression of P19 from *Cymbidium ringspot virus* can also inhibit RNA silencing by sequestering siRNAs and can lead to increased transgene expression[215].

1.3.5 Insulators

A challenge with the concept of a "part" is that it may perform differently in different positions in the chromosome or when surrounded by different parts, known as genetic context effects[216, 217]. This has led to new classes of genetic parts that reduce these effects, known as insulators[218-222].

Boundary elements can be used to insulate transgenes from local chromosome positional effects. Boundary elements are believed to act in two ways; first by interfering with enhancer activity by physically separating enhancers from the promoter, and second by protecting a flanked transgene from position-dependent silencing. Matrix attachment regions (MARs) are sequences of DNA where the nuclear matrix attaches, resulting in structural organization of the chromatin. MARs from tobacco, petunia, and *Arabidopsis*, as well as chicken, have been used to insulate transgenes in planta[223-226]. Usage of these elements led to increased transgene expression and/or the reduction of the expression variation between plants. Several boundary elements have been screened in *Arabidopsis*, each showing an increase in transgene expression levels yet only one, AtS/MAR10, also showed reduced variability between lines[227]. High-throughput open chromatin mapping technologies, such as Hi-C, can aid in identifying boundary elements in *Arabidopsis*, rice, cotton, and maize, that may be of use for mining or guiding the design of additional insulators[62, 228].

1.4 Programming plants to respond to their environment

Plants continually survey their environment and dynamically turn genes on and off to respond to shifting conditions or as a defense against threats[229]. Growth also requires the control of gene expression, where cells differentiate at different states to sprout, grow new leaves, control root morphology, and flower[230]. Chemical communication signals are transmitted between organelles within plant cells, between cells and tissues, and exogenously with colonizing bacteria and mycorrhizal fungi as well as between entire plants[9, 231]. These processes are controlled by a large regulatory network involving thousands of biochemical interactions that sense changes, process this information, and calculate the correct pattern of gene expression. Within minutes, this network responds to microbial infections or tissue damage done by insects

or animals[232, 233]. Within hours, it responds to shifts in soil conditions, such as nutrient availability, salinity, water content, and microbial signals, to shift carbon storage or remodel root morphology[234, 235]. Changes in temperature, humidity, and length-of-day lead to seasonal and cyclical responses[229].

This dynamic control of gene expression is in stark contrast to the control implemented over transgenes that are typically constitutively expressed, in other words, always on[236, 237]. It would be better to produce some traits only when they are needed, for example, turning on a cold resistance trait when the temperature is low[238]. This can be done by building synthetic sensors, encoded in the genome, that respond to environmental stimuli by turning an "output" promoter on or off. The output of a sensor can be used to control a gene, or it can be connected to a synthetic genetic circuit that implements a signal processing function to implement a desired response[238-241]. The circuit output can be connected to the control of transgenes or native genes. Many such sensors and circuits have been developed in bacteria, yeast, and mammalian cells, but have been slower to be adopted in plants[242-244]. One can imagine a future where plants are programmed to continuously survey their environment, process the information to determine which traits need to be activated, and implement the dynamics of a response.

1.4.1 Mapping and manipulating natural regulatory networks

Natural plant regulatory networks are large and complex and are only beginning to be understood as a system [245]. For example, in *Arabidopsis* there are >1,700 transcription factors (TFs) that collectively have 1.6 million interactions with genes, 14 enhancers that impact chromosome accessibility, 20 MAPKs, 100 miRNAs, and 1,000 peptides that target 600 receptors[246-251]. These regulators have an enormous impact on native gene expression; for example, the central glucose regulator KIN11 affects over 1,000 genes[229, 235, 252]. Within these networks, there are motifs that can produce logic operations, feedback loops, cascades, and cross regulation[234, 253].

Chromatin immunoprecipitation sequencing (CHIP-seq) is a powerful tool used to identify DNA binding sites of TFs and has recently been used to identify over 2 million binding sites for 104 maize leaf TFs, a median of ~16,000 binding sites per TF[254]. DNA affinity purification sequencing (DAP-seq) can take this one step further by showing how the binding of TFs is impacted by epigenetic modifications, such as methylation which affected 76% of surveyed TFs in *Arabidopsis*[255]. Alternatively, assays such as yeast one hybrid screens, provide a promoter-based approach to identifying potential TFs that bind to a promoter of interest and has been used to create gene regulatory networks involved with cell wall differentiation[256], defense response[257], and cell reprograming[258]. To accelerate the mapping of GRN in non-model plant species, such as Eucalyptus, groups are collecting parts to ease the move into DAP-seq and Y1H pipelines[259].

Additional computational approaches can be used to create GRNs, such as a Boolean network to show the crosstalk between phytohormones involved with induced system resistance (ISR) in Arabidopsis colonized with *Paraburkholderia phytofirmans* then infected with *Pseudomonas syringae*[260]. A bioinformatics approach was used to identify over 400 known and novel sequence motifs conserved in pathogen-coregulated genes and used to create 25 experimentally validated pathogen-inducible synthetic promoters in a parsley protoplast system and *N. benthamiana* [261]. Using 63 representative plant species, functional regulatory maps have been created with over 20 million transcription factor binding sites identified and 2 million functional interactions for over 20,000 transcription factors [262]. Computational tools have been developed to take a large number of maize RNA-seq data sets and identify potentially novel regulatory interactions[263].

With a better understanding of natural regulatory networks, targeted perturbations can be made to improve traits. For example, in *Arabidopsis* the overexpression of specific transcription factors increases cold and salt tolerance[264] and knocking another out improves disease resistance[229]. For metabolic engineering applications, natural transcription factors can be overexpressed to turn on biosynthetic pathways that are silent or to increase titers. For example, expression of the snapdragon transcription factors *Del* and *Ros1* from a tomato fruit-specific promoter resulted in purple tomatoes with high levels of anthocyanins[265]. Roots can be directed to grow deeper with more root hairs, thereby improving nutrient uptake and drought resistance, by overexpressing native transcription factors in rice, plumb and *Arabidopsis*[266, 267]. Gene regulatory networks also control the ratios of secondary metabolites, which can be perturbed to alter taste and aroma[268]. For example, apricot RNA-seq data from three different stages of ripening were used to generate a proposed pathway for flavor compound biosynthesis containing ~1200 differentially expressed genes with 16 potential regulators[268]. Finally, targeted RNA silencing can be an effective tool to regulate expression of native transcription factors[269].

Simple mutations in transcription factors can also be used to modulate regulatory networks in plants. The yield of rice has been improved by expressing a repressor mutated to be phytohormone-insensitive to divert biomass to grain[270, 271]. In rice, the balance between yield and immunity can be tuned by the phosphorylation state of a transcription factor, IPA1 [272]. Inducible overexpression of the same transcription factor can improve both rice yield and disease resistance [273]. Some targeted manipulations to what were believed to be well characterized pathways, such as climacteric fruit ripening, have emphasized the vast redundancy of native gene regulatory networks and supports the desire for synthetic regulation[274].

1.4.2 Commonly-used synthetic regulators

Synthetic regulators provide a platform for manipulating genes separate from the complexity of native GRNs. Common synthetic transcriptional regulators are often constructed from a small toolbox of

proven parts and have two primary components; a sequence specific DNA-binding domain to localize the regulator and a functional domain to manipulate transcription.

TFs can be fused to repressing domains, such the SRDX sequence (aka EAR-motif, ERF-associated amphiphilic repression) which is believed to disrupt the formation of the transcription complex or recruit native co-repressors *in planta* to restrict transcription[275-277]. Due to its small size of 12-aa, it is unlikely to inhibit protein folding and has been used in fusion with native TFs[276], heterologous zinc finger TFs to improve salt tolerance in ryegrass[278], as well as targeted with TALEs in *Arabidopsis*[277] and dCas9 in *Nicotiana benthamiana* leaves[279]. Additional repression motifs have been identified and can be localized to targeted promoters to repress transcription[280, 281].

The most common activating domain used is from the Herpes Simplex Virus VP16 protein [282] which interacts with chromatin remodelers and with general transcriptional machinery such as TFIID and TFIIH[277, 283]. The VP16 domain has been fused to ZFPs, TALEs, and dCas9 domains to create synthetic transcriptional activators in a wide range of plants. Other activating domains originating from plant transcription factors have been identified based on sequence similarity with the VP16 domain and a few were shown to have improved activation in tobacco after rational mutations were made to better match the VP16 sequences known to recruit TFIIH[284].

There are a few commonly ported protein regulators from other organisms that have been shown to function in plants. A fusion of the Tet repressor from the Tn10 transposable element to VP16 in stable tobacco can drive expression from a synthetic promoter containing tet operators and TATA box[285]. Yeast transcription factors representative of different families (MADS, Homeobox, GATA, and bZIP), including Gal4, have also been shown to function as activators or repressors in *Arabidopsis* and tobacco and the functionality could be modulated with the addition of VP16 or C1 activating domains[165, 286]. More complex chimeric regulators have also been made. For example, the XVE system uses the DNA-binding domain of bacterial LexA (X), the VP16 activating domain (V), and the regulatory region of the human estrogen receptor (E) to create an estradiol-activated regulator that has been used in Arabidopsis, tobacco, and rice among others[287].

Orthogonal regulators can also be built using dCas9 by designing sets of sgRNAs, which has been shown in yeast, bacteria, and plants[279, 288-290]. Similarly, TALEs and ZFPs can be designed to bind DNA sequences that are not present in the plant genome[291-294]. One caveat with the transfer of DNA binding domains from other kingdoms, is that they would typically not be allowed access to the plant nucleus without incorporation of a nuclear localization signal (NLS). One of the most common NLSs used is simian virus 40 (SV40) motif which is highly effective in a wide range of plants [295]. To adjust the dynamics of synthetic regulators a destabilizing domain can also be incorporated, such as amino acid residues 21 to 30

of the KIP-RELATED PROTEIN1(KRP1) which has been shown to increase turnover of synthetic regulators in Arabidopsis [296]

Riboswitches are fusions of ligand-binding RNA aptamers and self-cleaving ribozymes that can be used as cis-regulatory elements at the mRNA level. Typically, riboswitches are placed in the 3' UTR and upon binding to the aptamer-targeted ligand the self-cleavage of the mRNA results in rapid degradation of the transcript. Aptamers can be selected to bind to nearly any desired target using systemic evolution of ligands by exponential enrichment (SELEX)[297, 298]. However, the only riboswitch to date to be used to control the expression of a nuclear gene in plants is the highly optimized theophylline-responsive riboswitch, which contains the self-cleaving hammerhead ribozyme from *Shistosomoa mansoni*, and when placed in the 3' UTR of a target gene, regulates gene expression in a reversible and theophylline dose dependent manner[299].

While not yet demonstrated as a synthetic regulator in plants, the control of chromatin remodeling factors could provide an additional tool to gene regulation. In yeast, the recruitment of the chromatin remodeling repression domain Mxi1 by fusion to dCas9 results in strong targeted transcriptional repression[288]. It is known that plants naturally use chromatin structure to regulate genes – with genes contained in highly packed regions of the chromatin being inaccessible to transcription factors. In Arabidopsis, a protein called MONOPTEROS responds to auxin by recruiting chromatin remodelers which increase the DNA accessibility of key flower primordium initiation regulators – in essence providing a switchable method to regulate access to large parts of the genome[300].

1.4.3 Genetically-encoded sensors in plants

A sensor receives an environmental or cellular signal and controls an "output" in response. Here, we focus on transcriptional sensors, where the output is a promoter because these can be easily connected to transgenes to control their response, or a factor (*e.g.*, dCas9) to regulate native genes[216]. Sensors can range from individual promoters known to be turned on under the desired conditions to completely synthetic *de novo* designed regulators that bind to a small molecule of interest. The number of such sensors for plants has been growing, with examples that respond to environmental conditions (*e.g.*, drought, high temperature), agrochemical inducers, phytohormones, cell-cell communication signals, and synthetic small molecules (*e.g.*, fentanyl, TNT)[301, 302]. Most of the work done to date involves moving a single sensor into one plant line as a proof-of-principle, but one can imagine a future where plants contain an array of synthetic sensors that provide a myriad of situational awareness to the plant.

1.4.4 Condition-responsive plant promoters and transcription factors

The simplest sensors are based on a native promoter that responds to the desired conditions. How the promoter does this does not have to be well-understood. Common ways to identify such promoters are to use microarrays or RNA-seq to find promoters that turn on in response to a particular stimulus. This strategy has been used to find promoters that respond to nutrient deficiencies (phosphorus, sulfur, magnesium, etc.), heavy metal contamination (copper, cadmium, nickel, etc.), and stress (heat, cold, drought, etc.)[303-308].

The lack of understanding around regulatory inputs of native promoters makes it possible that the promoter may be affected by additional condition-dependent signals[305]. One way to reduce the number of inputs is to design a synthetic promoter with well-defined binding sites for native transcription factors. This requires a clean core promoter architecture into which operators for the transcription factor can be placed, a process aided by computational tools and databases of plant regulators[121, 309-315]. *Cis*-motifs have been identified that strongly association with heat shock, light, development, wounding, pathogen attack, sugar sensing, reactive oxygen species, and cold stress, and incorporation of these *cis*-elements can be used to create synthetic promoters that respond to different stimuli. Similarly, synthetic promoters can be made tissue-specific by inserting *cis*-elements that are bind transcription factors only found in certain tissues[316, 317].

Synthetic promoters can also be designed to bind synthetic transcription factors in order to minimize the reliance of native transcription factors which may themselves be embedded in complex, uncharacterized networks that respond to unknown signals. Synthetic transcription factors can be built using a ligand-biding domain from a plant regulatory protein with a DNA-binding domain containing activation or repression domains discussed in the previous sections. This approach was taken to build a set of synthetic phytohormone sensors that respond with high specificity to jasmonic acid, gibberellic acid, and auxin (so-called hormone activated dCas9-based repressors, or HACRs)[318]. Plant domains were used that recruit proteolytic machinery when bound to these molecules. These domains were fused to dCas9, targeted to the desired promoter, and a repression domain so that when the ligand binds, the fusion protein is degraded and the promoter turns on.

A red light sensor was built in tobacco protoplasts using a split transcription factor[319]. Each portion of a split transcription factor is unable to regulate transcription on its own and only when co-localized, often by the binding of a targeted molecule, is the full functional transcription factor assembled[320, 321]. In the presence of red light, an activating domain bound to phytochrome B changes conformation to bind to its cognate interacting factor tethered to a DNA binding domain and initiates transcription. When exposed to far-red light the conformation of phytochrome B switches back to its OFF state, unable to bind to the interacting factor and resulting in deactivation of transcription[322].

1.4.5 Porting sensors from other species

Regulators from other species, particular fungi or bacteria, are more likely to be orthogonal to the native plant gene regulatory networks. When a ligand binds to a regulatory protein, this typically leads to a conformational change that causes the protein to bind or unbind an operator sequence in a promoter[323]. These regulators can be ported to plants by fusing them to a nuclear localization signal and placing their operators in a plant minimal promoter. The binding event can be used to activate or repress a promoter by fusing the regulator to VD16 or SRDX domain, respectively[277]. This approach has been applied to build antibiotic (anhydrotetracycline) and ethanol sensors based on bacterial TetR-family repressors[285, 324, 325]. Ligand-binding domains sourced from other eukaryotes can also be used to build plant sensors, including quinic acid (mold), glucocorticoid (rat) and estrogen (human) sensors[326-328].

To build a sensor for oxygen orthogonal to native sensors in *Arabidopsis*, a split transcription factor and an oxygen sensitive enzyme from humans was used. One half of the split TF contained a yeast DNA binding domain (Gal4) and a human pVHL (von Hippel-Lindau tumor suppressor) while the other contained a yeast activating domain (Gal4) fused to a human HIF (hypoxia-inducible factor). In the presence of normal oxygen conditions, the enzyme PHD3 (human prolyl-hydroxylase) facilitates the hydroxylation of the HIF domain which heterodimerizes with pVHL to form a fully functional transcriptional activator – thus in the presence of oxygen driving expression of genes downstream of the Gal4 DNA binding sites[296]. This sensor shows that mammalian genes can be used in plants to create a sensor for hypoxic conditions, which could result from floods or waterlogged conditions. When flooded, plant tissue is starved of oxygen and being able to sense the concentration of available oxygen can be used to turn on necessary stress response genes.

Agriculturally-relevant chemicals could be used to induce gene expression in the field. A sensor for the commercially available insecticide methoxyfenozide was built by fusing the ligand-binding domain from spruce budworm ecdysone receptor (EcR) to a yeast DNA-binding domain (Gal4 or LexA) and VP16 activating domain[326-329]. The sensitivity of the sensor was further improved by using a two-hybrid approach; a GAL4 DNA-binding domain with *Choristoneura fumiferana* ecdysone receptor, and VP16 domain with *Locust migratoria* retinoid X receptor[330]. These sensors have been used in *Arabidopsis* and tobacco as well as corn and soybean protoplasts. A sensor for the non-steroidal ecdysone agonist tebufenozide was built using the ligand binding domain from *Heliothis virescens* ecdysone receptor, the VP16 domain, and the DNA binding domain of the mammalian glucocorticoid receptor and can induce transgene activity up to 420-fold in tobacco [331]. Agrochemical sensors have a better chance of being useful in the field than several other inducible systems but have not yet been verified[332].

1.4.6 De novo sensor design

It is challenging to create a sensor that responds to a molecule for which there is no natural binding domain. However, as the structures of more plant receptors are being solved, it has become possible to use this information to mutate them to bind to new ligands. Computational protein design or directed evolution can be used to alter the amino acids of a ligand binding pocket to make it accept a new ligand[333, 334].

Using the design software Rosetta, the binding pocket of a eukaryotic regulator was mutated to bind the pharmaceuticals digoxin, fentanyl, and progesterone[335-337]. Alternatively, rational mutations in the binding domain can be used to create a protein that is stable only when bound to a desired molecule - an approach that was used to create a sensor for digoxin in *Arabidopsis* by fusing the ligand-dependent stabilization domain to Gal4 and VP16[336]. A bump-and-hole strategy, a computational method to expand a binding pocket to bind to a larger molecule, was used to rationally design an orthogonal auxin receptor to bind to synthetic auxin-like molecules to probe auxin regulated networks *in planta*[338].

PYR1 is an abscisic acid (ABA) binding enzyme which when induced can increase drought tolerance in plants by, among other things, controlling guard cell aperture. Once structural information about the binding pocket of PYR1 was elucidated, a library of single amino acid mutations to each of the 25 residues could be screened against a panel of synthetic agrochemicals using a yeast two-hybrid assay. Hits were further improved using saturation mutagenesis of the remaining binding pocket residues and functional selections until a mutant, PYR1-MANDI was found to have nanomolar affinity to the fungicide mandipropamid. Arabidopsis and tomato expressing PYR1-MANDI show increased drought tolerance when induced by mandipropamid[339].

Signaling networks often rely on phosphorelays to integrate stimuli or transmit a signal to the nucleus. The simplest phosphorelay motifs in plants and bacteria are two-component systems, where a sensor histidine kinase phosphorylates a response regulator that binds to DNA[340]. A sensor for the explosive TNT (trinitrotoluene) was constructed in *Arabidopsis* by fusing a computationally-designed ligand-binding domain to the prokaryotic histidine kinase PhoR, which phosphorylates PhoB[341]. When fused to VP16, PhoB activated a plant promoter and TNT concentrations as low as 10 pM could be detected. The response, however, was limited by high background activity, which is likely due to cross reactions with endogenous histidine kinases.

An alternative approach to using proteins to bind to targeted molecules uses RNA, such as a riboswitch fused to the mRNA encoding a regulator that acts on a promoter[342]. When a small molecule binds to the riboswitch, there is a conformational change that alters the regulator expression level, thus changing the promoter activity. The plant-derived TPP (thiamine pyrophosphate) riboswitch controls 3'-UTR splicing and it can be moved to other genes to make them downregulated by TPP allowing for post-transcriptional regulation making it the first demonstration of riboswitch regulation of nuclear-encoded genes *in planta*[299, 343, 344]. A synthetic theophylline riboswitch activates a hammerhead ribozyme that cleaves

the mRNA to which it is fused[299] and provides another lever to control endogenous and recombinant genes in *Arabidopsis*[299, 343-345].

1.4.7 Computing in planta with synthetic circuits

Genetic circuits perform computational operations to process information from sensors[242]. The inputs and outputs of transcriptional circuits are defined to both be promoters, easing the connection to upstream sensors and downstream plant response. Signal processing can be applied to a single sensor, including noise filters, thresholding, inversion, or making the response more graded or switch-like. Multiple sensors can be integrated by digital or analogue logic, such as AND or NOR gates. Responding to more sensor activities increases the specificity of the response; in other words, more signals can be used to define an environmental condition. Dynamic circuits can delay a response after the signal is received or implement a pulse (adaptation) or oscillatory response. Memory permanently records a transient signal. To date, most genetic circuit research has been done in prokaryotes and the eukaryotic work has focused on yeast and mammalian cells [216, 346-349]. One of the challenges with building circuits in plants is that, in practice, properly connecting regulators by balancing their expression levels requires many trial-and-error attempts, which is slowed by long design cycles. Despite this, circuits are slowly being developed for plants so that they can perform computing operations in the field. Circuits may also aid in the control of development to express genes only at the correct stage or implemented needed dynamics. For example, it has been postulated that a synthetic oscillatory circuit controlling auxin could be used to induce lateral root formation at regular intervals[350].

Signal processing of single input sensors are some of the simplest circuits. In Arabidopsis, a TALE-VP64 (4 copies of the VP16 domain) based system was used to create a positive feedback loop by using the same native promoter to drive the synthetic TF as the targeted native gene. This system showed robust activation of certain targeted genes (while also showing synthetic transcription factor induced gene silencing for more rigorously regulated genes and miRNAs)[351]. A signal amplifier can be made from the simplified version of the quinic acid gene cluster from *Neurospora crassa*, referred to as the Q-system[352]. It contains a transcriptional activator (QF), a cognate DNA binding site (QUAS), and a repressor (QS) which when expressed blocks the activity of the activator. The Q-system can amplify a weak input by controlling the expression of QF which will amplify expression of a gene of interest downstream from a QUAS based promoter[326]. The Q-system has also been used to control the expression of multiple genes with a single input, which can be viewed as an alternative form of signal amplification[326].

There are a couple of strategies for performing logic operations on multiple inputs inside a plant cell. One approach is to place multiple operators in a single promoter; for example, if two regulators can both turn on a promoter, this will produce an OR function[68, 353]. Different logic operations can be realized

by changing whether the transcription factors binding the single promoter result in activation or repression. A library of synthetic transcriptional regulators from prokaryotes were shown to effectively activate or repress targeted gene expression downstream of synthetic promoters with minimal crosstalk – providing the necessary components to create several genetic circuits in planta. These parts were used to construct simple Boolean logic gates (OR, NOR, A NIMPLY B, and B NIMPLY A) where the output could be controlled by a single synthetic promoter[354]. In addition, using tissue specific promoters in Arabidopsis roots to drive the inputs of these circuits, 7 of the 9 tested logic gates correctly created the expected unique spatial expression patterns[354].

The affinity between the DNA binding domain and the cis-regulatory element can also affect the function of the synthetic promoter. A library of synthetic promoters based on mutated yeast Gal4 binding sites and random cis-elements upstream of a minimal promoters in tobacco showed expected trends in promoter strength given the strength of the cis-elements incorporated[165]. Synthetic promoters were also designed to bind three synthetic transcription factors – two which would activate transcription and one that would repress – though the promoters were not tested with simultaneous expression of the transcription factors[165]. Combining more than a few operators in a single promoter is technically challenging and requires mutational tuning to create the desired signal integration. An alternative is to "layer" gates by connecting them together to produce a more complex digital response. Layering of synthetic transcription factors his has been done recently for two-input circuits to create A IMPLY B, B IMPLY A, AND, and NAND gates in N. benthamiana transient assays[354].

Two-input Boolean logic gates were made in Arabidopsis protoplasts and plants using two unique recombinases – Flp and B3 – and recombination sites flanking a strong terminator and/or the promoter or a reporter gene[355]. These circuits were used to control expression of a reporter in a targeted cell type by using a cell type specific promoter as one input and small molecule inducer as the other. The major limitation to recombinase-based circuits is that they are not reversible, yet this could also be viewed as a perk since they essential act as permanent memory.

The simultaneous binding of an activator and a repressor can lead to a multitude of logic operations. Paradoxical components can provide robust input-output relationships where the output depends on the input signal and not on the concentration of any of the proteins involved in the circuit [356]. These types of systems are natively present in plants - the PPDK system in the C4 pathway contains a bifunctional enzyme RP which makes the output of the pathway robust to variations in substrate and protein levels[357]. In addition, a circuit which combines a negative autoregulation loop with relatively weak positive autoregulation shows increased response dynamics in *E coli*[358] but these systems have yet to be built synthetically *in planta*.

Logic operations can also be accomplished using split transcription factors which are only active when both components are present, essentially making an AND gate. A split-TALE system was developed

by separating the DNA binding domain from the activating domain of a TALE and fusing both to interacting protein domains which constitutively interact with each other[359]. This system resulted in targeted gene expression of levels around 30% of those induced by a 35S promoter. A Flp recombinase could be split to enforce the simultaneous presence of two inputs – each controlling the expression of one half of the recombinase – adding the restriction that both inputs have to be present at the same time to recombinase based circuits[355]. Alternatively, placing a protease cleavage site between moieties allowed for targeted restoration of transcription when the appropriate protease was simultaneously expressed in tobacco transient assays[360]. Incorporating multiple cleavage sites and expressing their cognate proteases allowed for the restoration of transcriptional activity following AND and OR logic[360].

Memory can convert a transient signal into a permanent response. So-called "volatile" memory requires the constant usage of energy and resources to maintain the state. For example, bistable switches, which can be formed by positive feedback loops or cross-repression, require the constant expression of a regulator[361, 362]. However, to date there are no published reports of successful "volatile" bistable toggle switches in plants despite being theorized in several papers[363]. Permanent memory does not require constant energy to maintain the state. This can be achieved by using enzymes that modify the state of the DNA, for example inverting a region where each orientation corresponds to a different state[364, 365]. Flipping a DNA sequence can orient a promoter or terminator to turn expression on or off. Such a memory switch has been built in tobacco using a phage integrase to invert DNA containing a promoter[366]. The integrase was placed under the control of an estradiol sensor and when plants are exposed transiently to this molecule, expression continued permanently.

The timing of desired gene expression can be tuned using native epigenetic pathways. Flower development involves a cascade of transcription factors, the expression of which can be repressed by presence of epigenetic markers, for example trimethylation is established and maintained by Polycomb group proteins which bind upstream of the transcription factor KNU. The displacement of polycomb group proteins allows for the removal of the epigenetic modifications over time due to cell division and ultimately expression of KNU approximately 2 days later. A synthetic epigenetic timer circuit was built by designing a promoter which contains binding elements for native polycomb group proteins as well as binding sites for a chimeric LacI-TAL protein. When the chimeric transcription factor is expressed, it competitively binds with polycomb group proteins, allows for the removal of the epigenetic modifications, and finally expression of the downstream gene with a 2 day delay[367].

1.4.8 Connecting synthetic circuits to plant phenotypes

The outputs of sensors and circuits can be used to control plant phenotype. The simplest approach is to simply place transgenes under the control of the output promoter. Using them to control a more

complex trait encoded by native genes in the genome, for example to alter a developmental pathway or control flux through a metabolic pathway, can be more difficult as it requires a means to use an output promoter to drive the up- and down-regulation of those genes. To this end, there are a number of technologies to control expression at the transcriptional, translational, or post-translational levels.

The output promoter can be used to control the expression of a natural transcription factor, thus interfacing the synthetic network with the natural one[368]. The regulatory protein can be mutated to eliminate any unwanted natural signals. For example, mutants to orthologs of the Arabidopsis Gibberellin Insensitive (GAI) in *Arabidopsis*, rice, wheat, and maize make them less sensitive to the phytohormone gibberellin and show dwarf phenotype with increased yield[270]. There are several examples where a transcription factor was placed under the control of a sensor to induce a phenotype at a desired time. To time rice flowering to optimize yield or react to inclement weather, a repressing transcription factor was constitutively expressed and then a second anti-repressor was placed under the control of sensors that respond to the agrochemicals Routine and Orzyemate[272, 369-371]. Similarly, the growth of petunia can be stunted at desired times by using a dexamethasone sensor to control the expression of a transcription factor[372]. Finally, BUFFER gates with different strength activators controlling the expression of a mutated developmental regulator IAA14 modulated lateral root branch phenotypes in Arabidopsis[354].

The output of a circuit can be connected to the transcriptional control of native genes by repurposing the DNA-targeting tools from genome editing. TALEs, ZFPs, and dCas9 can be designed to repress endogenous genes by fusing them to activation or repression domains[373-375]. In *Brassica*, for example, ZFPs have been used to reduce saturated fatty acid content of canola oil by elevating the expression of two canola β-ketoacyl-ACP synthase IIs (*KASII*)[376]. Alterations to native gene networks in response to phytohormones via the incorporation of synthetic transcription factors can be used to manipulate plant morphology. An auxin-sensitive HACR targeting the promoter of an auxin transporter (PIN1) was used to alter the strength by which auxin regulates its own abundance in the shoot (reducing the gain of a positive feedback loop) ultimately reducing shoot branching by nearly half in *Arabidopsis*[318]. Simultaneously targeting multiple genes could redirect carbon flux, mediate complex phenotypes (*e.g.*, lignin biosynthesis), or coordinate stress responses. As a proof-of-principle, TALEs were fused to VP64 activating domains to simultaneously activate three *Arabidopsis* genes[377]. It is relatively easy to have dCas9 target multiple genes because of the simplicity of designing sgRNA; in bacteria, up to 22 sgRNAs have been used to regulate 13 genes simultaneously[352, 374]. A multiplexed CRISPR system, named CRISPR-Act2.0, uses a modified gRNA scaffold with bacteriophage aptamers that recruit four extra VP64 activating domains in tandem with dCas9:VP64. This system was shown to robustly activate three genes simultaneously in rice protoplasts[377]. Using dCas9 to simultaneously localize VP64 activation domains and TAL activation domains results in even stronger transcriptional activation of multiple genes in *Arabdopsis*[375] .

Circuits can also control expression of native genes by expressing factors that target their mRNA. Many crop traits have been improved through use of siRNAs, including enhancement of plant disease and pest resistance, alternation of plant architecture, and removal of toxic compounds[115, 266, 378-381]. To control root growth in *Medicago*, a dexamethasone sensor has been used to drive the expression of hairpin RNA, which is processed by DICER-LIKE endonucleases into siRNAs, to knockdown a gene required for fatty acid synthesis and thus critical for root survival[382]. Gentian flower color has also been modified by downregulating enzymes involved with anthocyanin and flavonoid modification pathways using hairpin RNA[383]. Another means to regulate expression at the mRNA level is to transcribe antisense RNA that complements the target sequence and blocks translation or results in transcript degradation[384]. To make it more digestible to livestock, a commercialized alfalfa (HarvXtra) was engineered to constitutively transcribing antisense RNA to knockdown a lignin biosynthesis enzyme in vascular tissue[385, 386]. This can be made more effective by transcribing multiple antisense RNAs targeting multiple biosynthetic enzymes in the lignin biosynthesis pathway and has been used to reduce lignin levels 20-30% in tobacco[387]. Finally, translation in tobacco and potato plastids can be controlled by regulation the expression of PPR protein, which is localized to the plastid and binds target RNA sequences[224, 388, 389].

Endonucleases can also be used to regulate expression at the mRNA level. The Csy4 endonuclease, sourced from bacterial CRISPR arrays, targets a 28 nt sequence inserted in the 5'-UTR and this has been shown to lead to inducible mRNA degradation in tobacco and rice[59]. Orthologs to Csy4 have also been identified and validated in planta[59]. dCas13a endonuclease can be directed to nearly any target mRNA with a sgRNA, removing the constraint of inserting target sequences to the 5'-UTR. It has been used in rice protoplasts and tobacco to knockdown targeted mRNAs, including to block RNA viruses[390, 391]. Expression of CRISPR/Cas13a can target and cleave viral RNA providing resistance to the host plant[390, 392].

Controlling the alternative splicing of precursor mRNAs can be used to regulate gene expression at the posttranscriptional level. A suicide exon P5SM (plant 5s rRNA mimic) from Oryza sativa can be placed within the ORF of a targeted gene which introduces a stop codon, prevents proper function, and targets the transcript for degradation. When expressed, the ribosomal protein L5 from Oryza sativa will splice out P5SM and increase transgene expression nearly 100-fold[393]. This strategy has been used to create a chemical-inducible "hypersensitive response" system by tightly controlling the expression of an effector protein in tobacco with non-detectable background expression[394].

Epigenetic modifications can also be used as an output from a genetic circuit in plants. DNA methylation is an epigenetic modification commonly involved in gene silencing. A DNA methyltransferase can be fused to zinc fingers to target promoters for methylation, thus silencing downstream gene expression[395]. In addition, a human demethylase can also be targeted to a desired promoter using a zinc finger, resulting in upregulation of a desired downstream gene[211]. Site specificity can be further improved

by using a methyltransferase:dCas9 fusion with multiple gRNAs targeting the desired promoter for methylation[213]. DNA methylation patterns modified using synthetic methyltransferases / demethylases in plants can result in heritable epialleles without changing the DNA sequence and can be used to control phenotypes such as flowering time[211, 213, 395].

1.5 Accelerating plant engineering

The slow growth of plants leads to tediously slow design, build, and test cycles[396]. Constructing new crop cultivars and evaluating them in the field can take years – for maize it used to take up to 11-13 years from the time of initial crosses to release on the market. Even making a stable laboratory strain of Arabidopsis can take several months. Due to low throughput of plant transformation, only a small number of transformants can be generated, which often do not produce the desired trait and errors are often not discovered until plants are regenerated and analyzed. This could be addressed by developing a method to prototype circuit designs before having to make a complete plant, for example by using transient assays, protoplasts, single-cell algae, cyanobacteria or cell-free systems[155, 397-399].

Transient testing of constructs is possible with agroinfiltration, but the results don't always provide an accurate representation of how things will function in a stable line due to variations in expression levels and co-suppression in RNA silencing. Using protoplasts could also rapidly accelerate the process of quantitative characterization of synthetic parts in plants. However, there can be a very high degree of batch variation in the preparation of the protoplasts [400]. Regardless, the use of transient assays and other imperfect yet high-throughput assays still result in fewer design cycles needed to optimize the system in whole plants.

Currently, *Zea mays* (maize), *Triticum aestivum* (wheat), *Solanum lycopersicum* (tomato) are among the few model crops used in plant research. However, their large and complex nuclear genome, slow life cycles, and often non-trivial transformation and regeneration methods pose engineering challenges. To that end, liverwort (*Marchantia polymorpha*)[401], green alga (*Chlamydomonas reinhardtii*)[402], Fast-Flowering Mini-Maize[403], *Brachypodium distachyon*[403], and *Seteria viridis*[404] have seen greater use as more tractable experimental chassis. For example, *Marchantia polymorpha* is a simple liverwort with a dominant haploid gametophytic lifestyle, easy propagation, and a smaller genome size with less genome redundancy than higher plants, making it an attractive model organism for plant biology [405, 406]. Its genome was sequenced and used to identify putative genetic parts such as promoters and TF binding elements [405]. *Marchantia* also has effective nuclear and chloroplast transformation methods, standardized vectors, antibiotic resistance genes, and fluorescent proteins, and is amenable to high throughput screening[406]. Exploration of these new model species can accelerate engineering efforts in more recalcitrant crops.

Protoplasts are cell-wall-free plant cells that can be transformed and sorted in an automated fashion to help the selection and analysis of nuclear transformed cells [407, 408 409]. Protoplasts have been used to probe

endogenous gene expression, identify candidate genes in important metabolic pathways [410], and quantitatively characterize genetic parts in different plant species such as *Arabidopsis* and sorghum[400 411]. Protoplasts can be sorted using fluorescence-activated cell sorting, for example to screen libraries of CRISPR-Cas9 based prime-editing of rice and wheat cells [156, 157]. Whole plants can be regenerated from isolated protoplasts, such as the prime-edited rice cells in the previous example which had a ~20% regeneration rate[156]. Protoplasts are also amenable to use with robotic systems where protoplast production, transformation, and analysis can all be automated for higher-throughput screens [407, 408].

The use of cell-free systems offers the possibility to further simplify and accelerate the testing of genetic constructs in plants. Wheat germ cell-free systems can express properly folded eukaryotic protein complexes and membrane proteins for high-throughput screening of genes [412, 413]. Automated liquid handling machines can couple DNA assembly and cell-free prototyping platform based on *E. coli* and wheat germ lysates to characterize and assign functions to several plant genes [414]. A BY-2 tobacco cell-free system that contains microsomal vesicles can promote proper eukaryotic protein folding, formation of disulfide bonds, glycosylation, and co-translational integration of membrane proteins [415 416].

Cell-free systems from purified chloroplast provide the transcription and translation machinery of the subcellular compartment and could enable an *in vitro* milieu for genetic-prototyping [417, 418]. Chloroplast extracts share some similarities to *E. coli* cell-free systems, in terms of prototyping genetic parts that bear similarities to the prokaryotic ones, like the plastid-encoded polymerase (PEP) promoters [419]. However, some prominent differences were also observed [420] suggesting that the utilization of the chloroplast cell-free systems be better suited for plastid genetic prototyping.

1.6 Plant-microbe communication channels

A key requirement for a holistic agricultural system synthetic biology approach is the ability to coordinate function across multiple organisms. Signaling molecules which are stable extracellularly and can trigger a transcriptional response upon detection provide a means of linking functions spread between multiple cells which may be physically separated. An ideal signaling system would be easily scalable, modular, and have minimal crosstalk with endogenous signaling systems. However, there are several examples of natural plant-microbe communication molecules that provide a starting point for future optimization. Endogenous signaling molecules such as flavonoids [421], AHLs [422], volatile organic compounds (VOCs), phytohormones [423], DKPs, and oxylipins have been used to engineer transkingdom communication in controlled environments. In these systems, typically either the biosynthesis of the molecule is placed under synthetic regulation, or the transcriptional regulator which binds to the molecule is engineered.

Plants nurture communities of bacteria and fungi in the rhizosphere by excreting carbon sources (sugars, organic acids, amino acids, etc.) and signaling compounds from their roots[424]. Engineering microbes to respond to plants exudates is a valuable way to only activate microbes when a specific need occurs or when proximal to a plant. In some cases, root exudates may be an indirect indication of plant status. For example, plant nitrogen deficiency decreases amino acid exudation in maize[425], phosphorus deficiency increases organic acid (citrate, malate, and succinate) exudation[426], and iron deficiency leads to coumarin exudation[426].

Activation of microbial sensors by exudates when inoculated onto plants has been demonstrated. In early studies of exudate spatiotemporal patterns, simple microbial biosensors were constructed in plant-associated bacteria *Pantoea agglomerans* by fusing native sucrose-, fructose-, tryptophan-, or galactose-responsive promoters to reporter genes[427-430]. A genetic sensor based on the native galactose/galactoside-induced *melA* promoter in *Sinorhizobium meliloti* could sense galactosides, sugars found in root exudates, up to 200 μm from the root surface[430]. More recently, a study screened promoters from the genome of the legume symbiont *Rhizobium leguminosarum* to identify those that respond to root exudates, from which sensors were constructed for erythritol, myo-inositol, carboxylates (formate, malonate, tartrate), amino acids (phenylalanine, GABA) and flavonoids (hesperetin)[431]. These sensors were inoculated onto pea roots to study the spatial and temporal changes in nodule root exudation but also provides evidence for using plant exudates to control gene expression in nearby soil bacteria[432].

Sensors can also be constructed by introducing heterologous regulators and promoters. To our knowledge, sensors have been built for the following known plant root exudates using the listed regulatory proteins: organic acids – cumate (CymR)[433], vanillate (VanR, QacR)[434-436], benzoate (BenM)[437], *p*-coumarate (PadR)[438, 439], protocatechuic acid (PcaV)[440], tartrate (TtdR)[437], succinate (DcuRS)[441], malate (MalKZ)[442], fumarate(DcuSZ)[443], and glutarate(CsiR)[444]; sugars – arabinose (araC)[434, 435, 437], xylose(XylR)[445], cellobiose (CelR)[446], galactose (GalR)[446], ribose (RbsR)[446], fructose (ScrR)[446], gentiobiose (LacI mutant)[447], and fucose (LacI mutant)[447]; flavonoids – naringenin (FdeR, TtgR)[434, 435, 448, 449], kaempferol (QdoR)[449], and quercetin (QdoR)[449]; amino acids – lysine (LysG)[450], arginine (ArgP)[450], tryptophan (TrpR)[451], β-alanine (OapR)[437], phenylalanine (PhhR)[437], GABA (GabR)[437], tyrosine (HpdA)[437], aspartate (Tar-EnvZ/OmpR)[452], glutamate (DegS-EnvZ/OmpR)[453]; and miscellaneous metabolites choline (BetI)[435], stilbenes (saro_0803)[454], and xanthine (XdhR)[437]. A sensor for the phytohormone cytokinin *trans*-zeatin was constructed by fusing the *Arabidopsis* AHK4 sensory domain to an engineered PhoQP two component system in *E. coli*[455]. The engineered system was orthogonal to native microbial signaling and responsive to 1μM *trans*-zeatin [455]. While to date, none of these heterologous sensors have been tested in conjunction with plants, these sensors may exhibit greater orthogonality to native microbial pathways and thus may enable tighter coupling of the desired output to the plant.

Flavonoids, despite being used in endogenous signaling, are a large and diverse family of molecules that can be engineered to create novel orthogonal signal-receiver pairs. On the transmitting end, our understanding of the enzymes responsible for flavonoid modifications allows for the biosynthesis of unnatural derivatives [456]. On the receiving end, rational design or directed evolution of the ligand binding pocket of lysR family transcriptional regulators could enable binding to unnatural flavonoids[421].

Microbes have also been engineered to sense several naturally plant-produced volatiles. VOCs are widespread in long-distance intra- and inter-species signaling cascades due to their physicochemical properties. Blends of VOCs released from plants can induce antibiotic resistance, virulence, and motility genes in neighboring bacteria or act as antifungals, repellents or attractants to different fungi, parasitoids, and entomopathogens. For example, VOC production from broad beans (*Vicia faba*) in response to aphid [457] feeding can recruit parasitic wasps (*Aphidius ervi*) which naturally kill the aphids. However, aphids which carry a particular endosymbiont (*Hamiltonella defensa*) have an increased survivability due to the symbionts ability to reduce and alter the blend of herbivore-induced plant VOCs [458]. Just as the endosymbiont *H. defensa* naturally alters the emitted blend of VOCs, it can be beneficial to engineer synthesis of a unique VOC fingerprint from a crop or microbe in a field to attract beneficial organisms such as pollinators or repel pathogenic ones. A wheat line engineered to constitutively produce an aphid alarm pheromone repelled cereal aphids and increased foraging of a parasitic wasp in a controlled environment[459]. However, this line showed no effect in field trials likely due to the constitutive expression suggesting the need for additional regulation, perhaps in the form of genetic circuits. Salicylic acid is a major plant hormone best known for its role in the plant defense response, but also is involved in plant growth, development, and abiotic stress tolerance[460]. The transcriptional activator NahR, which recognizes salicylic acid, has been ported to build sensors in *E. coli* [435, 437] and in soil bacteria *Cupriavidus necator* [437], *Pseudomonas putida* [437], *Rhizobium* sp. IRBG79[461], and *Azorhizobium caulinodans*[461]. Isoprene is an abundant plant-produced volatile [462] that may be involved in thermoregulation and photoprotection and possibly yet unknown functions[463]. *E. coli* and *Pseudomonas putida* engineered to express high levels of the regulator TbuT were responsive to isoprene production in culture, but this sensor was also activated by phenol, benzene, and toluene and thus may not be specific enough for agricultural purposes[464].

Small RNAs (sRNAs) in plants are essential for regulating host immunity by controlling the stability and translation of target mRNAs. sRNAs can travel from cell to cell, root to shoot, as well as between organisms where they can induce gene silencing in the counterparty [465, 466]. Due to their ability to suppress virulence, the direct application of sRNAs to crops has gained traction as an alternative approach to disease control [467]. Alternatively, transgenic plants can produce artificial sRNA targeted towards pathogen virulence genes[468]. In addition, the transkingdom exchange of sRNAs via extracellular vesicles could be exploited by engineering plant associated microbes to exchange sRNAs towards targeted genes of

interest in a host plant. Both of these techniques have broad applicability due to the evolutionary conserved sRNA trafficking mechanism and could be further engineered to control the expression of any gene of interest in either party.

One drawback of the sensors previously described in this section is that they involve ubiquitous plant exudates, so non-specific activation is a concern. This can be solved by engineering plants to secrete a molecule that can activate a microbial sensor. Phloroglucinol, a common precursor to antimicrobials natively made by several plant growth promoting *Pseudomonas*, can be produced in *Arabidopsis* by expressing a type III polyketide synthase from *Pseudomonas fluorescens* Pf-5 in the plastid[469]. Synthetic sensors have been designed which enable *Pseudomonas protegens* Pf-5, *Rhizobium sp.* IRBG74, and *E. coli* to respond to this molecule[434, 435, 470]. Plants have also been engineered to produce opines, carbon substrates that are produced in crown galls when a pathogen hijacks host plant metabolism, and rhizopines, inositol derivatives that perform a similar function to opines but are produced by some *Rhizobium*. Barley has been engineered to excrete the rhizopine *scyllo*-inosamine by constitutively expressing 2 genes from bacteria, *Rhizobium leguminosarum* and *Sinorhizobium meliloti*, and it can be detected from roots by a genetic sensor carried in *Rhizobium leguminosarum*[471]. Nitrogen fixation has been placed under the control of synthetic opine (octopine and nopaline) sensors in *Azorhizobium caulinodans*[434]. Several opine biosynthesis and sensor genes have been identified from *Agrobacterium* and could be used for additional channels of communication between plants and microbes. It should be noted, however, that the optimal molecule for synthetic plant-to-microbe communication channels should not be found naturally in the rhizosphere, and while phloroglucinols, opines, and rhizopines are only naturally produced by a few species, they are still not ideal. Ultimately, the ideal molecules would be compounds not found in the natural world.

1.7 Synthetic symbioses

The symbiotic relationships between plants and microbes plays a vital role in overall plant health. These microbes primarily reside in the rhizosphere, the region impacted by plant exudates, as well as on (epiphytes) and inside (endophytes) the root tissue itself. Several studies have shown that plants of difference species, geographical locations, climates, and land management practices show distinct rhizosphere microbiomes [472]. In addition, plant exudates are dynamic and change depending on species, developmental stage, root traits, environmental conditions, nutrition, and soil type [472]. Taken together, it becomes evident that plants are actively selecting for beneficial symbionts within the rhizosphere. These beneficial bacteria can provide plant growth promoting traits such as auxin production [473], ethylene regulation [474], phytopathogen suppression [475], and nutrient acquisition [476]. Engineering synthetic symbiosis

is a promising strategy to provide these benefits, or rationally design completely new relationships, to crops of interest.

One of the original proposals for maintaining plant-microbe symbiosis is to engineer the root exudates of the plant – a plan based on the opine theory [477]. Not only can opines act as signaling molecules but they can act as the sole carbon and nitrogen source of bacteria with the appropriate catabolic genes. *Lotus corniculatus* genetically engineered to produce opines was shown to shift the microbiome composition towards microbes capable of catabolizing the specific opines being produced [478]. Furthermore, *Pseudomonas fluorescens* engineered with a 19kb fragment that confers opine catabolism from *Agrobacterium* reached a higher population density than *P. fluorescens* without the opine catabolism genes when co-inoculated at equal concentrations in otherwise sterile soil of opine-producing transgenic tobacco [479].

1.8 Conclusion

As a field, agriculture stands to benefit greatly from the advances in genetic engineering currently being made. The genetic parts and regulators available for use in plants has grown rapidly in recent years and will lead to crops with improved yields and higher tolerance to a changing climate. We are beginning to have options in regards to where certain traits should be placed within a plant with emerging tools for plastid engineering. Despite these advances, genetic engineering *in planta* is much more complex than what is capable in microbes which have relatively simple and rapid lifecycles. To best take advantage of our engineering capabilities in microbes, they should be viewed as an integral part of the broader agricultural system – essentially an external plastid. Attaining this view requires forms of communication between plants and microbes in order to link the entire system. Looking to nature as a guide, there are several examples of communication molecules which may be repurposed in an engineered system by building new sensors or placing their biosynthesis under synthetic regulation. However, the search for ideal signals that are completely orthogonal remains an outstanding challenge in engineering a holistic agricultural system.

2 Control of nitrogen fixation in bacteria that associate with cereals

This chapter is composed of select sections of an article titled, "Control of nitrogen fixation in bacteria that associate with cereals", which was published in Nature Microbiology in February 2020 and written by Min-Hyung Ryu and Christopher Voigt. I was a third author on the paper and helped build and test a few of the sensors included in the final section, "Control of nitrogen fixation with agriculturally relevant sensors." That portion from the publication, along with a minimally altered introduction, discussion, and relevant methods are included below.

2.1 Introduction

In agriculture, nitrogen is a limiting nutrient that needs to be added as fertilizer to those crops that cannot produce it on their own, including the cereals rice, corn, and wheat. In contrast, legumes are able to obtain nitrogen from the atmosphere using nitrogen-fixing bacteria that reside in root nodules. However, the majority of the world's calories are from cereals; thus, it has been a longstanding problem in genetic engineering to transfer this ability to these crops[480, 481]. This would reduce the need for nitrogenous fertilizer and the economic, environmental, and energy burdens that it brings[482]. One solution is to engineer the bacteria that associate with cereals to fix nitrogen, whether they be in the soil, on the root surface (epiphytes) or living inside of the roots (endophytes)[483].

Some rhizobia isolated from legume root nodules are also cereal endophytes[484-487], however most are unable to fix nitrogen under free-living conditions (outside of the nodule) [488, 489]. There have been reports of cereal yield improvements due to these bacteria, including a 20% increase for rice by *Rhizobium* sp. IRBG74, but this is likely due to other growth-promoting mechanisms, such as improved nutrient uptake or root formation[488, 490-514]. *Azorhizobium caulinodurans* ORS571 is exceptional because it is able to fix nitrogen in both aerobic free-living and symbiotic states, has been shown to be a rice and wheat endophyte, and does not rely on plant metabolites to produce functional nitrogenase[507, 515-519]. However, when *Rhizobium* or *Azorhizobium* species are living in cereal roots, there is low nitrogenase expression and $^{15}N_2$ transfer rates suggest any reported uptake is due to bacterial death[488, 490-509]. To date, it has not been shown that a *Rhizobium* strain can be engineered to fix nitrogen under free-living conditions when it does not do so naturally.

Several bacterial species are used as cereal seed inoculants that either fix nitrogen naturally or are potential hosts into which the capability could be transferred. The non-host-specific endophyte *Pseudomonas stutzeri* and epiphyte *Klebsiella oxytoca* colonize rice and wheat and are used to improve growth[520-526]. The epiphyte *Pseudomonas protegens* Pf-5 produces a suite of antibiotics and is used as a commercial biocontrol seed inoculant for maize and rice, but cannot fix nitrogen[527, 528].

The nitrogen fixation (*nif*) genes are organized as genomic clusters, ranging from a 10.5kb single operon in *Paenibacillus* to 64kb divided amongst three genomic locations in *A. caulinodans*. Conserved genes include those encoding the nitrogenase enzyme (*nifHDK*), FeMoCo biosynthesis, and chaperones[529-534]. Species that can fix nitrogen under more conditions tend to have larger gene clusters that include environment-specific paralogues, alternative electron transport routes, and oxygen protective mechanisms[529-540]. Often, the functions of many genes in the larger clusters are unknown.

There is evolutionary evidence for the lateral transfer of *nif* clusters between species[541, 542]. However, achieving such a transfer via genetic engineering poses a challenge as many things can go awry, including differences in regulation, missing genes, and the intracellular environment[489, 535, 543-545]. In 1972, the first engineered transfer from *K. oxytoca* to *E. coli* was reported and subsequently clusters from other species have been functionally moved to this host[545-552]. The cluster from *P. stutzeri* has been moved to *P. protegens* Pf-5 and other pseudomonads[540, 553, 554]. The small *Paenibacillus polymyxa* WLY78 cluster has been transferred to soil isolates of *Bacilli*[549].

Nitrogenase is under stringent control because it is oxygen sensitive and energetically expensive: it can make up 20% of the cell mass and each ammonia requires ~40 ATP[555, 556]. It is also irreversibly deactivated by oxygen. Across species, transcription of *nif* genes is strongly repressed by fixed nitrogen (ammonia) and oxygen with these signals converging on the NifA regulatory protein that works in concert with the sigma factor RpoN[534, 555-558]. Diverse, species-specific, and often poorly understood signals control

these regulators, including plant-produced chemicals, ATP, reducing power, temperature, and carbon sources[545, 555-557, 559-563]. Those bacteria that can fix nitrogen in a wider range of environmental conditions tend to be controlled by more complex regulatory networks[489, 555-557].

When a *nif* cluster is transferred from one species to another, it either preserves its regulation by environmental stimuli or has an unregulated constitutive phenotype[535, 545, 549, 554]. Maintaining the native regulation, notably ammonium repression, limits their use in agriculture because such levels are likely to fluctuate according to soil types[564], irrigation[565], and fertilization[566]. Nitrogen-fixing diazotrophs have been engineered to reduce ammonia sensitivity by disrupting NifL[567, 568] or mutating NifA[569-572] and placing the entire cluster under the control of T7 RNA polymerase (RNAP)[549-552, 571-573]. Constitutive expression of nitrogenase is also undesirable as it imparts a fitness burden on the cells[495, 574]. For example, when the *nif* cluster from *P. stutzeri* A1501 was transferred to *P. protegens* Pf-5, this was reported to result in sufficient ammonia production to support maize and wheat growth, but the bacterial population quickly declined after a month when competing with other species in soil[554, 575]. Constitutive activity is detrimental even before the bacteria are introduced to the soil, impacting production, formulation, and long-term storage[528, 576]. Therefore, uncontrolled nitrogenase production could lead to more expensive production, shorter shelf life, and more in-field variability[491, 561, 577, 578].

In this manuscript, we present several strategies to implement control over nitrogen fixation in bacteria that live on or inside the roots of cereals. Two approaches are taken: either the native regulation is replaced or a *nif* cluster is transferred from another species and placed under inducible control. For *A. caulinodurans*, ammonium-independent control is achieved using a sensor to drive the co-expression of a NifA mutant and RpoN in a Δ*nifA* strain. *Rhizobium* sp. IRBG74 can be engineered to express functional nitrogenase under free living conditions either by transferring a native *nif* cluster from *Rhodobacter* or a refactored cluster from *Klebsiella*. Multiple approaches enable *P. protegens* Pf-5 to express functional nitrogenase, of which the transfer of the *nif* cluster from *P. stutzeri* A1501 and *Azotobacter vinelandii* DJ yields higher activity and oxygen tolerance.

Based on these systems, we demonstrate the control of nitrogenase using genetically-encoded sensors that respond to signals of relevance to cereal agriculture. There are many potential signals that could be used to keep nitrogenase off at undesirable times and then switch it on when fixed nitrogen can be delivered to the target plant. In the soil, sensors could respond to biocontrol agents or components of added fertilizer and other treatments (*e.g.*, DAPG: 2,4-diacetylphloroglucinol)[579-581]. Proximity to the plant could be detected by root exudates[582], including sugars, hormones, flavonoids, and antimicrobials (*e.g.*, arabinose, salicylic acid, naringenin and vanillic acid)[488, 518, 519, 583-591]. Sensors that respond to these signals have been previously shown to turn on in bacteria that are close to the roots[592]. We also demonstrate control in response to non-cereal exudates that are involved in remodeling the microbial community (*e.g.*, cuminic acid)[593, 594].

Other bacteria colonizing the cereal root surface, including exogenously added plant growth promoting bacteria, release chemicals that could also be used to detect proximity (*e.g.*, 3,4-dihydroxybenzoic acid, 3OC6HSL and 3OC14HSL)[595-601]. Plants could also be genetically modified to excrete a chemical signal that is then sensed by the engineered microbes[602, 603]. To this end, pathways have been previously introduced into plants that lead to the secretion of opines, phloroglucinol, and rhizopene[478, 604-607]. As a proof-of-concept for the response to a synthetic communication signal, we demonstrate the control of nitrogenase with opines (nopaline and octopine). Collectively, this work presents an unprecedented side-by-side comparison of control strategies for the transfer and manipulation of nitrogenase pathways in diverse bacteria.

2.2 Results

2.2.1 Control of nitrogen fixation with agriculturally-relevant sensors

The careful design and characterization of the controller has the benefit of simplifying the process by which different synthetic sensors are used to induce nitrogenase expression. By knowing the dynamic range required to go from inactive to active nitrogenase, one can quantitatively select sensors that produce a compatible response. This allows different environmental signals – or combinations of signals using genetic logic circuits – to be used to control expression. To demonstrate this, we selected 11 synthetic sensors that respond to a variety of chemical signals of relevance to the rhizosphere and demonstrate that these can be used to create inducible nitrogenase in our engineered strains of *E. coli* (carrying the refactored v2.1 *nif*), *R.* sp. IRBG74 (carrying the refactored v3.2 *nif*), *P protegens* Pf-5 (carrying the inducible *A. vinelandii nif*), and *A. caulinodans* (inducible *nifA/rpoN*) (Figure 2.1).

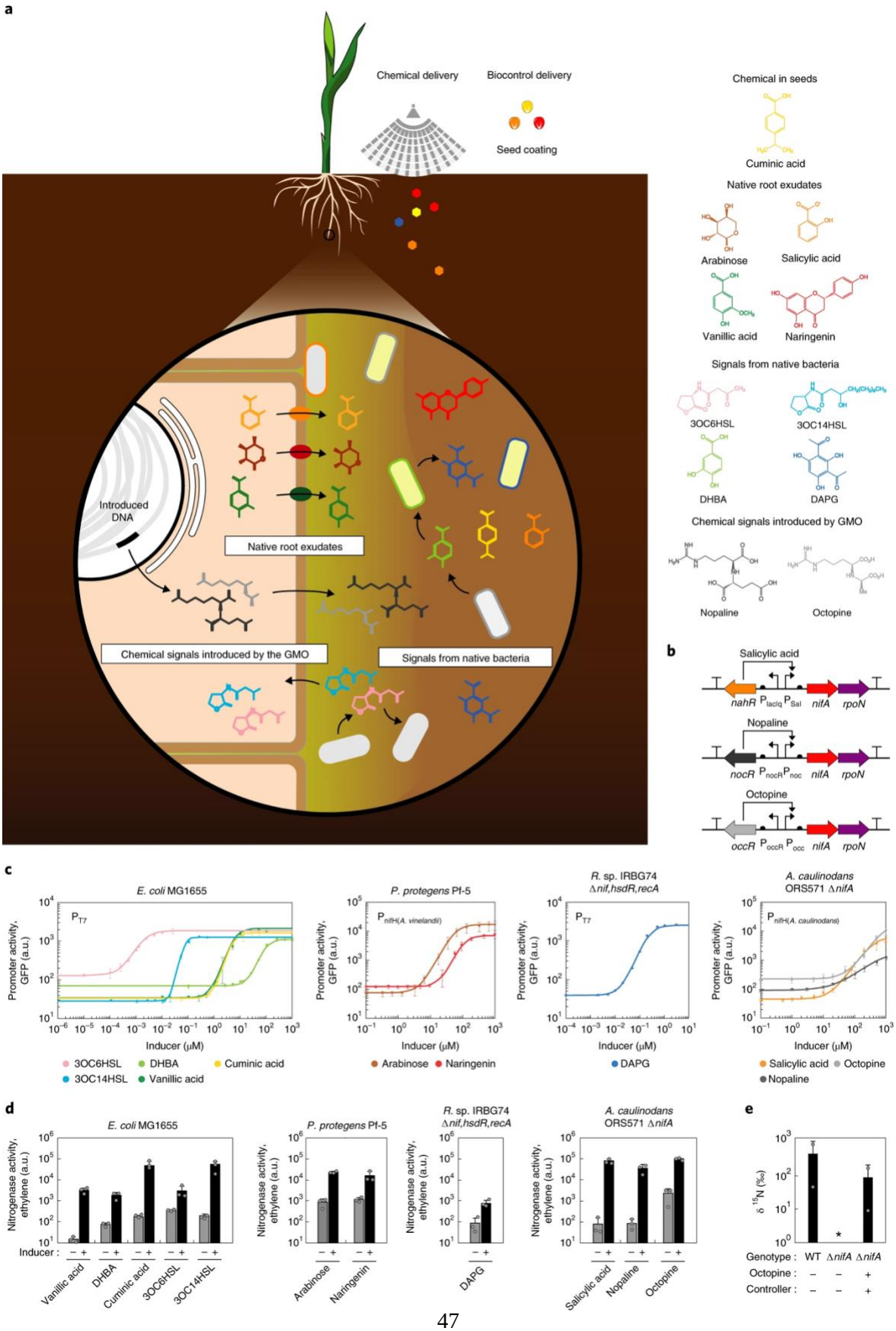
a
Chemical delivery
Biocontrol delivery
Seed coating
Introduced DNA
Native root exudates
Chemical signals introduced by the GMO
Signals from native bacteria
Chemical in seeds
Cuminic acid
Native root exudates
Arabinose
Salicylic acid
Vanillic acid
Naringenin
Signals from native bacteria
3OC6HSL
3OC14HSL
DHBA
DAPG
Chemical signals introduced by GMO
Nopaline
Octopine
b
Salicylic acid
nahR PlaclqPSal nifA rpoN
Nopaline
nocR PnocRPnoc nifA rpoN
Octopine
occR PoccR Pocc nifA rpoN
c
E. coli MG1655
P. protegens Pf-5
R. sp. IRBG74 Δnif,hsdR,recA
A. caulinodans ORS571 ΔnifA
Promoter activity, GFP (a.u.)
Inducer (μM)
3OC6HSL
3OC14HSL
DHBA
Vanillic acid
Cuminic acid
Arabinose
Naringenin
DAPG
Salicylic acid
Octopine
Nopaline
d
E. coli MG1655
P. protegens Pf-5
R. sp. IRBG74 Δnif,hsdR,recA
A. caulinodans ORS571 ΔnifA
Nitrogenase activity, ethylene (a.u.)
Inducer :
Vanillic acid
DHBA
Cuminic acid
3OC6HSL
3OC14HSL
Arabinose
Naringenin
DAPG
Salicylic acid
Nopaline
Octopine
e
δ 15N (‰)
Genotype : WT ΔnifA ΔnifA
Octopine : – – +
Controller : – – +

Figure 2.1 Control of nitrogen fixation with agriculturally relevant sensors.
a, Schematic showing the origins of the chemicals. Introduced DNA, genetic modification of the plant to produce nopaline and octopine. GMO, genetically modified organism. b, Genetic sensors built for controlling nitrogenase activity in A. caulinodans. c, Response functions of the sensors. Either the sensor expresses T7 RNAP, which then activates PT7, or it expresses nifA/(rpoN) and activates the nifH promoter (the species origin is indicated in parentheses). d, The nitrogenase activity was measured in the presence or absence of inducer (see Methods). The refactored K. oxytoca nif clusters v2.1 and v3.2 were used in E. coli MG1655 and R. sp. IRBG74, respectively. The inducible A. vinelandii nif cluster was used in P. protegens Pf-5. The controller containing nifA/rpoN was used in A. caulinodans ΔnifA. e, Incorporation of 15N into cell biomass. Nitrogen fixation in the wild-type A. caulinodans ORS571, A. caulinodans ΔnifA and A. caulinodans ΔnifA carrying the controller, in which nifA/rpoN is inducible by octopine (+1 mM octopine), was traced using 15N2 and analysed using isotope-ratio mass spectrometry (see Methods). The asterisk indicates 15N incorporation at levels below the detection limit. The inducers were used at the following concentrations: 50 μM vanillic acid, 500 μM DHBA, 50 μM cuminic acid, 25 nM 3OC6HSL, 500 nM 3OC14HSL, 33 μM arabinose, 100 μM naringenin, 100 nM DAPG, 200 μM salicylic acid, 1 mM nopaline and 1 mM octopine. The error bars represent the s.d. from three (d) or two (e) independent experiments performed on different days.

The roles of the chemical signals in the rhizosphere are shown in Figure 2.1a. Cuminic acid is present in plant seeds [593, 594, 608] and functions as a fungicide[609]. Natural root exudates may include sugars, amino acids, organic acids, phenolic compounds, phytohormones, and flavonoids[610, 611]. These represent potential signals to control nitrogenase production close to the root surface. Cereals have been shown to release arabinose, vanilic acid, and salicylic acid[591, 612-617]. In addition, salicylic acid regulates the plant innate immune response and the impact of its exogenous addition to cereals has been studied[618]. Naringenin is a common precursor for many flavonoids and improves endophytic root colonization when applied to rice and wheat[517, 518]. Genistein, a product from naringenin catalyzed by the isoflavone synthase, is released from maize roots[619]. A quorum sensing mimic released by rice can regulate the 3OC6HSL receptor protein LuxR, which has been visualized using *E. coli* biosensor strains[620].

Bacteria either native to the rhizome or added as biocontrol agents introduced as a spray inoculant or seed coating produce chemical signatures. Inoculation of cereals with root colonizing *Pseudomonas* strains that produce DAPG elicit protection against fungal pathogens[581, 621]. Many bacteria produce quorum molecules, such as N-acyl homoserine lactones, as a means of communication and plants can respond to these signals[596, 597, 602, 622-625]. The bacterium *Sinorhizobium meliloti* produces 3OC14HSL, which enhances *Medicago* nodulation and has been shown to induce systemic resistance in cereals[623, 626]. DHBA can be produced by root colonizing bacteria to increase iron solubility[600, 601] and play a role as a chemoattractant for *Agrobacterium* and *Rhizobium*[627].

Sensors for these chemicals were constructed based on the controllers for each species. For *E. coli* MG1655, we have previously constructed a strain that contains 12 optimized sensors, carried in the genome, that respond to various small molecules ("Marionette")[628]. The response functions of these sensors were characterized in standard units, making it simple to identify those that can be connected to nitrogenase

expression without further tuning. Marionette contains sensors for vanillic acid, DHBA, cuminic acid, 3OC6HSL, and 3OC14HSL. For each sensor, the output promoter was transcriptionally fused to T7 RNAP and the response of the responsive promoter (P_{T7}) was measured as a function of inducer concentration (Figure 2.1c). Then, the v2.1 refactored *nif* cluster was introduced and nitrogenase activity was measured in the presence and absence of inducer (Figure 2.1d).

The inducible systems constructed for *P. protegens* Pf-5 that respond to arabinose and naringenin were used to drive NifA expression for the control of the *A. vinelandii nif* cluster (Figure 2.3). The induction of the *nifH* promoter by these sensors was first confirmed using a reporter (Figure 2.1c). When this is replaced with the *nif* gene cluster, this results in an inducible response of nitrogenase activity (Figure 2.1d). The best nitrogenase activity in *R.* sp. IRBG74 is low; however, we sought to demonstrate that it could be placed under inducible control. The DAPG-inducible system developed for *R.* sp. IRBG74 was connected to the control of T7 RNAP and this produces a strong response from P_{T7} (Figure 2.1c). However, when used to drive the expression of the v3.2 refactored pathway, only a 9-fold induction is observed, consistent with the low nitrogenase activity observed in this strain (Figure 2.1d). Finally, the salicylic acid sensor designed for *Rhizobium* was used to control NifA (L94Q/D95Q)/RpoN expression in *A. caulinodans* (Figure 2.1b/c and Figure 2.2). This yielded a 1000-fold dynamic range of nitrogenase activity (Figure 2.1d).

Plants could be engineered to release an orthogonal chemical signal that could then be sensed by a corresponding engineered bacterium[483, 583]. This would have the benefit of only inducing nitrogenase in the presence of the engineered crop. Further, if the molecule is metabolizable by the engineered bacterium, it could serve as a mechanism around which a synthetic symbiosis could be designed, where the plant provides the carbon and the bacterium fixed nitrogen in an engineered relationship. To this end, legumes and Arabidopsis have been engineered to produce opines, including nopaline and octopine[478, 607]. We constructed sensors for these two opines for *A. caulinodans* based on the LysR-type transcriptional activators OccR (octopine) and NocR (nopaline) and their corresponding P_{occ} and P_{noc} promoters (Figure 2.1b and Figure 2.4). These sensors were connected to the expression of NifA(L94Q/D95Q)/RpoN and the response from P_{nifH} was measured using a fluorescent reporter. Both response functions had a large dynamic range (Figure 2.1c) and produced highly-inducible nitrogenase activity (Figure 2.1d). The nopaline sensor yielded a 412-fold dynamic range and the octopine sensor led to 40% higher nitrogenase activity than the wild-type.

2.3 Discussion

Towards designing a bacterium that can deliver fixed nitrogen to a cereal crop, this work provides a side-by-side comparison of diverse species, natural *nif* clusters, and engineering strategies. The goal was

to obtain inducible nitrogenase activity in a strain that can associate with cereals as an endophyte or epiphyte. To this end, we constructed ~100 strains involving the transfer of 10 natural *nif* clusters ranging in size from 10kb to 64kb to 16 diverse species of *Rhizobia*, *Azorhizobium*, *Pseudomonas*, and *E. coli*. Different approaches were taken to make these *nif* clusters inducible, from bioinformatics and protein engineering to complete genetic reconstruction from the ground-up (refactoring). In addition to the highest activity, it is important that nitrogen fixation be robust to the addition of nitrogenous fertilizer (ammonia) and microaerobic environments. Two lead candidates have emerged from this effort. Our most promising endophyte is a variant of *Azorhizobium* where *nifA* is knocked out of the genome and a *nifA* mutant and *rpoN* are complemented on a plasmid. For the epiphyte *P. protegens* Pf-5, the most versatile strain is based on the transfer of the *A. vinelandii nif* cluster and placement of *nifA* of *P. stutzeri* under inducible control. In both cases, nitrogenase activities were obtained that are nearly identical to wild-type *A. caulinodurans* and *P. stutzeri*, respectively. Neither showed significant repression by ammonia and optimal activity was obtained in 1% oxygen. Based on these strains, we demonstrate that nitrogenase can be placed under inducible control in response to cereal root exudates (arabinose, salicylic acid), phytohormones (naringenin) and putative signaling molecules that could be released by genetically modified plants (nopaline and octopine).

Because *R.* sp. IRBG74 can fix nitrogen in a legume nodule and also associates with rice, we put significant effort into engineering this strain to fix nitrogen when cereal-associated. This has proven to be surprisingly difficult. We first tried simply complementing *nifV*, as this is absent in *R.* sp. IRBG74 and produces a metabolite provided by the plant, but these attempts were unsuccessful. Then, it was found that all of the initial *nif* clusters transferred, some of which have high activity in *P. protegens* Pf-5 and *E. coli*, are non-functional in *R.* sp. IRBG74. This led to us trying clusters from alphaproteobacteria, one of which produced a very low level of activity that was dependent on the *nif* genes native to *R.* sp. IRBG74. The previously-published refactored gene clusters based on *Klebsiella nif* were attempted in *R.* sp. IRBG74 but these showed no activity. It was only after the construction of a new refactored cluster (v3.2) that we obtained activity under free-living conditions that was not dependent on the native *nif* genes. This allowed us to increase the expression levels, and we discovered an optimum beyond which activity was lost. To our knowledge, this is the first time that *nif* activity has been engineered in a *Rhizobium* under free-living conditions that could otherwise not perform this function. Note, however, that the activity is quite low, and our measurements are only based on acetylene reduction. Because we were not able to get higher activity, we opted to not pursue the further development of this strain.

This paper encompasses different degrees of *nif* pathway re-engineering to promote heterologous transfer. The most ambitious is the complete refactoring of all the *nif* genes and regulation, where all regulatory genetic parts are replaced, genes are recoded, operons are reorganized, and transcription is

performed by the orthogonal T7 RNAP. When this project was initiated, DNA synthesis was a novelty and a lack of DNA assembly methods made it difficult to make alternative designs. Further, the evaluation of performance relied on the overall nitrogenase activity, rather than an understanding of the underlying parts. As such, the first refactored pathway performed poorly[550]. In subsequent studies, better part libraries and DNA assembly and automation platforms[551, 629] enabled us to make many variants. Further, as the cost of RNA-seq declined, we were able to use to this to evaluate the performance of internal parts, such as promoters and terminators. This showed us that our first designs were effectively large single operons with little differential control over the transcription levels of individual genes. With these techniques, we were able to optimize the function of the refactored *nif* pathway and discovered that many of the underlying genetic structure was unnecessary to achieve high activities.

In the current work, we applied ribosome profiling, a new technique that enables the measurement of translational parts (*e.g.*, ribosome binding sites) and infer expression levels. Further, nitrogenase activity and the function of underlying parts were assessed as the clusters were moved between species. Interestingly, the native *Klebsiella nif* cluster could be transferred and it performed similarly but the refactored cluster yielded widely varying expression levels in the different hosts, sometimes leading to a total loss in activity. This could be recovered by maintaining the native operon structure in the refactored cluster, implying that it was not due to the synthetic sensors, T7 RNAP, or promoters/terminators. This is one of the hypothesized functions of operons[257, 630, 631]. Achieving this required maintaining the codon usage and translational coupling of the native cluster. However, this does not mean that it will not be possible to also encode this function synthetically. There have been computational advances that enable the calculation of RBSs internal to upstream genes when encoded on an operon[632]. If coupled with codon optimization algorithms, this would allow the design of *de novo* genetic parts that achieve a desired degree of translational coupling and expression level.

This work is the first step of a larger effort to build strains that can efficiently deliver fixed nitrogen to cereals[483, 583]. Here, we show the deregulation of *nif* clusters in *A. caulinodurans* and *P. protegens* Pf-5, enabling them to be placed under the control of cereal root exudates. This derepresses the pathway in the presence of exogenous nitrogenous fertilizer – critical for the use of the bacterium as part of an integrated agricultural solution. Further, these organisms retain the ability to fix nitrogen in microaerobic environments, thus avoiding the need for a root nodule that enforces strict anaerobiosis. The complete deregulation of the *nif* pathway makes the bacterium non-competitive in the soil and lost quickly, thus limiting its impact to particular phases of the growth cycle. Thus, we demonstrated that nitrogenase can be placed under the control of chemical root exudates. Fully realizing the goal of engineering microbial delivery to a cereal will require significant additional genetic engineering to maximize their ability to catabolize carbon sources from the plant and increase the flux of fixed nitrogen delivery by redirecting

metabolism, introducing transporters, and the optimization of electron transfer. An intriguing possibility is to also genetically engineer the plant to produce orthogonal carbon sources, such as opines or less common sugars, and then placing the corresponding catabolism pathways into the bacterium.

2.4 Materials and Methods

Only methods from Ryu et al. related to the included result section are below. For the complete detailed materials and methods see Ryu et al.[434].

2.4.1 Bacterial strains and growth media.

E. coli DH10-beta (New England Biolabs, MA, Cat# C3019) was used for cloning. *E. coli* K-12 MG1655 was used for the nitrogenase assay. *P. protegens* Pf-5 was obtained from the ATCC (BAA-477). *R.* sp. IRBG74 was obtained through the courtesy of Dr. Jean-Michel Ané. *P. stutzeri* A1501, *A. caulinodans* ORS571 and 12 rhizobia strains were obtained through the courtesy of Dr. Phil Poole. *A. vinelandii* DJ was obtained through the courtesy of Dr. John Peters. For rich media, LB medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl), LB-Lennox medium (10g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl), and TY medium (5 g/L tryptone, 3 g/L yeast extract, 0.87 g/L $CaCl_2 \cdot 2H_2O$) were used. For minimal media, BB medium (0.25 g/L $MgSO_4 \cdot 7H_2O$, 1 g/L NaCl, 0.1 g/L $CaCl_2 \cdot 2H_2O$, 2.9 mg/L $FeCl_3$, 0.25 mg/L $Na_2MoO_4 \cdot 2H_2O$, 1.32 g/L $NH_4CH_3CO_2$, 25g/L Na_2HPO_4, 3 g/L KH_2PO_4 pH 7.4), UMS medium[633] (0.5 g/L $MgSO_4 \cdot 7H_2O$, 0.2 g/L NaCl, 0.375 mg/L EDTA-Na_2, 0.16 $ZnSO_4 \cdot 7H_2O$, 0.2 mg/L $Na_2MoO_4 \cdot 2H_2O$, 0.25 mg/L H_3BO_3, 0.2 mg/L $MnSO_4 \cdot H_2O$, 0.02 mg/L $CuSO_4 \cdot 5H_2O$, 1 mg/L $CoCl_2 \cdot 6H_2O$, 75 mg/L $CaCl_2 \cdot 2H_2O$, 12 mg/L $FeSO_4 \cdot 7H_2O$, 1 mg/L thiamine hydrochloride 2 mg/L D-pantothenic acid hemicalcium salt, 0.1 mg/L biotin, 4 mg/L nicotinic acid, 87.4 mg/L K_2HPO, 4.19 g/L MOPS pH 7.0), and Burk medium[634] (0.2 g/L $MgSO_4 \cdot 7H_2O$, 73 mg/L $CaCl_2 \cdot 2H_2O$, 5.4 mg/L $FeCl_3 \cdot 6H_2O$, 4.2 mg/L $Na_2MoO_4 \cdot 2H_2O$, 0.2 g/L KH_2PO_4, 0.8 g/L K_2HPO_4 pH 7.4) were used. Antibiotics were used at the following concentrations (μg/mL): *E. coli* (kanamycin, 50; spectinomycin, 100; tetracycline, 15; gentamicin, 15). *P. protegens* Pf-5 (kanamycin, 30; tetracycline, 50; gentamicin, 15; carbenicillin, 50), *R.* sp. IRBG74 (neomycin, 150; gentamicin, 150; tetracycline, 10; nitrofurantoin, 10), and *A. caulinodans* (kanamycin, 30; gentamicin, 15; tetracycline, 10; nitrofurantoin, 10).

2.4.2 Strain construction.

A *sacB* markerless insertion method was utilized to allow deletions and replacements of a native locus with synthetic parts by homologous recombination. In order to increase transformation efficiency in *R.* sp. IRBG74, a type-I restriction-modification system was inactivated by deleting *hsdR*, which encodes a restriction enzyme for foreign DNA[635]. Two homology arms of ~500bp flanking the *hsdR* gene were

amplified by PCR, cloned and yielded a suicide plasmid pMR44. The suicide plasmid was mobilized into *R.* sp. IRBG74 by triparental mating. Single-crossover recombinants were selected for resistance to gentamicin and subsequently grown and plated on LB plates supplemented with 15% sucrose to induce deletion of the vector DNA part containing the counter selective marker *sacB* which converts sucrose into a toxic product (levan). Two native *nif* gene clusters encompassing *nifHDKENX* (genomic location 219,579-227,127) and *nifSW-fixABCX-nifAB-fdxN-nifTZ* (genomic location 234,635-234,802) of *R.* sp. IRBG74 were sequentially deleted using pMR45-46. To increase genetic stability[636] *recA* gene was deleted using the plasmid pMR47. The *R.* sp. IRBG74 Δ*nif, hsdR, recA* strain was the basis for all experiments unless indicated otherwise. Two homology arms of ~900bp flanking the *nifA* gene were amplified by PCR, cloned and yielded a suicide plasmid pMR47 to generate *nifA* deletion in *A. caulinodans* ORS571. The suicide plasmid pMR47 in *E. coli* was mobilized into *A. caulinodans* by triparental mating. Single-crossover recombinants were selected for resistance to gentamicin and subsequently grown and plated on TY plates supplemented with 15% sucrose to induce deletion of the vector DNA part. All markerless deletions were confirmed by gentamicin sensitivity and diagnostic PCR.

2.4.2 Plasmid system.

Plasmids with the pBBR1 origin were derived from pMQ131 and pMQ132[637]. Plasmids with the pRO1600 origin was derived from pMQ80[637]. Plasmids with the RK2 origin were derived from pJP2[638]. Plasmids with the RSF1010 origin were derived from pSEVA651[639]. Plasmids with the IncW origin were derived from pKT249[550].

2.4.6 Transformation.

Electroporation was used to transfer plasmids into *P. protegens* Pf-5. A single colony was inoculated in 4 mL of LB and grown for 16 h at 30°C with shaking at 250 rpm. The cell pellets were washed twice with 2 mL of 300 mM sucrose and dissolved in 100 μl of 300 mM sucrose at RT. A total of 50-100 ng DNA was electroporated and recovered in 1 mL of LB media for 1 h before plating on selective LB plates. Triparental mating was used to transfer DNA from *E. coli* to rhizobia. An aliquot of 40 μl of late-log phase (OD_{600} ~ 0.6) donor cells and 40 μl of late-log phage helper cells containing pRK7013 were mixed with 200 μl of late-log phase (OD_{600} ~ 0.8) recipient rhizobia cells and washed in 200 μl of TY medium. Mating was initiated by spotting 20 μl of the mixed cells on TY plates and incubated at 30°C for 6 h. The mating mixtures were plated on TY medium supplemented with nitrofurantoin to isolate rhizobia transconjugants.

2.4.7 Construction and characterization genetic parts for rhizobia

Genetic part libraries were built on the pBBR1-ori plasmid pMR1 using Gibson assembly (New England Biolabs, Cat#E2611). The fluorescence proteins, GFPmut3b and mRFP1 were used as reporters. The Anderson promoter library[640] on the BioBricks Registry were utilized for the characterization of constitutive promoters. To characterize inducible promoters, a regulator protein is constitutively expressed by the P_{lacIq} promoter, and GFP expression is driven by a cognate inducible promoter from the opposite direction, facilitating replacement of the reporter with gene of interest (e.g., T7 RNAP and *nifA*) and transfer of the controller unit across different plasmid backbones for diverse microbes. The following combinations of cognate regulators and inducible promoters were characterized: IPTG inducible LacI-$P_{A1lacO1}$, DAPG inducible PhlF-P_{Phl}, aTc inducible TetR-P_{Tet}, 3OC6HSL inducible LuxR-P_{Lux}, salicylic acid inducible NahR-P_{Sal}, and cuminic acid inducible CymR-P_{Cym} systems were optimized for *R.* sp. IRBG74 (Figure 2.2). Opine inducible OccR-P_{occ}, and nopaline inducible NocR-P_{noc} systems were optimized for *A. caulinodans* (Figure 2.4). For RBS characterization, an IPTG-inducible GFP expression plasmid pMR40 was used and GFP was expressed to the highest levels with 1 mM IPTG. RBS library for GFP was designed using the RBS Library Calculator[641] at the highest-resolution mode, and the 3' end of the 16S rRNA sequences were adjusted according to the species (3'-ACCTCCTTC-5' for *R.* sp. IRBG74). Terminators for T7 RNAP were characterized by placing a terminator between two fluorescence reporters expressed from a single T7 wild-type promoter located upstream of the first fluorescence protein GFP[173]. The expression of the two fluorescence proteins is enabled by the controller strain MR22 encoding the IPTG-inducible T7 RNAP system by 1 mM IPTG. The terminator strength (Ts) was determined by normalizing fluorescence levels of a terminator construct by a reference construct pMR67 where a 40 bp spacer was placed between the reporters. All genetic parts for *R.* sp. IRBG74 were characterized as follows. Single colonies were inoculated into 0.5 ml TY supplemented with antibiotics in 96-deepwell plates (USA Scientific, Cat#18962110) and grown overnight at 30°C, 900 rpm in a Multitron incubator (INFORS HT, MD). 1.5 µl of overnight cultures was diluted into 200 µl of TY with antibiotics and appropriate inducers in 96-well plates (Thermo Scientific, Cat# 12565215) and incubated for 7 h at 30°C, 1,000 rpm in an ELMI DTS-4 shaker (ELMI, CA). After growth, 8 µl of culture sample was diluted into 150 µl phosphate buffered saline solution (PBS) with 2 mg/mL kanamycin for flow cytometry analysis.

2.4.8 Construction and characterization genetic parts for P. protegens

Genetic part libraries were built on a pRO1600-ori plasmid pMR2 using Gibson assembly (New England Biolabs, Cat#E2611). The fluorescence proteins, GFPmut3b and mRFP1 were used as reporters. The Anderson promoter library[640] on the BioBricks Registry were utilized for the characterization of constitutive promoters. The following combinations of cognate regulators and inducible promoters were characterized: IPTG inducible LacI-P_{tac}, DAPG inducible PhlF-P_{Phl}, aTc inducible TetR-P_{Tet}, 3OC6HSL

inducible LuxR-P_{Lux}, arabinose inducible AraC-P_{BAD}, cuminic acid inducible CymR-P_{Cym}, and naringenin inducible FdeR-P_{Fde} were optimized (Figure 2.3). For RBS characterization, an arabinose-inducible GFP expression plasmid pMR66 was used and GFP was expressed with 7 μM arabinose. RBS library for GFP was designed using the RBS Library Calculator[641] at the highest-resolution mode, and the 3' end of the 16S rRNA sequences were adjusted according to the species (3'-ACCTCCTTA-5' for *P. protegens* Pf-5). The expression of the two fluorescence proteins is enabled by an IPTG-inducible T7 RNAP expression system of the controller strain MR7 by 0.5 mM IPTG. All genetic parts for *P. protegens* Pf-5 were characterized as follows. Single colonies were inoculated into 1 ml LB supplemented with antibiotics in 96-deepwell plates (USA Scientific, Cat#18962110) and grown overnight at 30°C, 900 rpm in a Multitron incubator (INFORS HT, MD). 0.5 μl of overnight cultures was diluted into 200 μl of LB with antibiotics and appropriate inducers in 96-well plates and incubated for 7 h at 30°C, 1,000 rpm in an ELMI DTS-4 shaker. After growth, 10 μl of culture sample was diluted into 150 μl PBS with 2 mg/mL kanamycin for flow cytometry analysis.

2.4.9 Genomic integration and characterization of controllers.

The mini-Tn7 insertion system[642] was used to introduce a controller into the genome of *P. protegens* Pf-5. The IPTG-inducible T7 RNAP expression system and a tetracycline resistant marker *tetA* was placed between two Tn7 ends (Tn7L and Tn7R), yielding the controller plasmid pMR86, which was introduced into *P. protegens* Pf-5 by double transformation with pTNS3[642] encoding the *TnsABCD* transposase. A genomically-integrated controller located 25 bp downstream of the stop codon of *glmS* was confirmed by PCR and sequencing. A markerless insertion method using homologous recombination was employed in *R.* sp. IRBG74 (described above). A controller encoding inducible T7 RNAP system flanked by two homology fragments that enables the replacement of *recA* was cloned into a suicide plasmid. These controller plasmids (IPTG-inducible, pMR82-84; DAPG-inducible, pMR85) in *E. coli* was mobilized into *R.* sp. IRBG74 MR18 (Δ*hsdR.* Δ*nif)* by triparental mating, generating the controller strains (MR19, 20, 21 and 22, respectively). The controller integration in the genome was confirmed by gentamicin sensitivity and diagnostic PCR. All controllers were characterized in a manner identical to that described in genetic part characterization.

2.4.10 Construction and characterization of Marionette-based controllers.

To regulate nitrogenase expression in the *E. coli* Marionette MG1655[628], the *yfp* in the 12 reporter plasmids was replaced with T7 RNAP while keeping other genetic parts (e.g., promoters and RBSs) unchanged. The reporter plasmid pMR121 in which *gfpmut3b* is fused to the T7 promoter variant P_2 was co-transformed to analyze the response functions of each of the 12 T7 RNAP controller plasmids. To

characterize controllers, single colonies were inoculated into 1 ml LB supplemented with antibiotics in 96-deepwell plates and grown overnight at 30°C, 900 rpm in a Multitron incubator. 0.5 µl of overnight cultures was diluted into 200 µl of LB with antibiotics and appropriate inducers in 96-well plates and incubated for 6 h at 30°C, 1,000 rpm in an ELMI DTS-4 shaker. After growth, 4 µl of culture sample was diluted into 150 µl PBS with 2 mg/mL kanamycin for flow cytometry analysis.

2.4.11 Flow cytometry.

Cultures with fluorescence proteins were analyzed by flow cytometry using a BD Biosciences LSRII Forterssa analyzer with a 488 nm laser and 510/20-nm band pass filter for GFP and a 561 nm laser and 610/20 nm band pass filter for mCherry and mRFP1. Cells were diluted into 96-well plates containing PBS supplemented with 2 mg/mL kanamycin after incubation. Cells were collected over 20,000 events which were gated using forward and side scatter to remove background events using FlowJo (TreeStar Inc., OR). The median fluorescence from cytometry histograms was calculated for all samples. The median autofluorescence was subtracted from the median fluorescence and reported as the fluorescence value in arbitrary unit (au).

2.4.12 Nitrogenase assay (E. coli and K. oxytoca).

Cultures were initiated by inoculating a single colony into 1 mL of LB supplemented with appropriate antibiotics in 96-deepwell plates (USA Scientific, Cat#18962110) and grown overnight at 30°C, 900 rpm in a Multitron incubator. 5 µl of overnight cultures was diluted into 500 µl of BB medium with 17.1 mM $NH_4CH_3CO_2$ and appropriate antibiotics in 96-deepwell and incubated for 24 h at 30°C, 900 rpm in a Multitron incubator. Cultures were diluted to an OD_{600} of 0.4 into 2 mL of BB medium supplemented with appropriate antibiotics, 1.43 mM serine to facilitate nitrogenase depression[643], and an inducer (if necessary) in 10 mL glass vials with PTFE-silicone septa screw caps (Supelco Analytical, Cat#SU860103). Headspace in the vials was replaced with 100% argon gas using a vacuum manifold. Acetylene freshly generated from CaC_2 in a Burris bottle was injected to 10% (vol/vol) into each culture vial to begin the reaction. The acetylene reduction was carried out for 20 h at 30°C with shaking at 250 rpm in an Innova 44 shaking incubator (New Brunswick) to prevent cell aggregations, followed by quenching via the addition of 0.5 mL of 4 M NaOH to each vial.

2.4.13 Nitrogenase assay (P. protegens Pf-5).

Cultures were initiated by inoculating a single colony into 1 mL of LB supplemented with appropriate antibiotics in 96-deepwell plates and grown overnight at 30°C, 900 rpm in a Multitron incubator. 5 µl of overnight cultures was diluted into 500 µl of BB medium with 17.1 mM $NH_4CH_3CO_2$ and appropriate antibiotics in 96-deepwell and incubated for 24 h at 30°C, 900 rpm in a Multitron incubator.

Cultures were diluted to an OD_{600} of 0.4 into 2 mL of BB medium supplemented with appropriate antibiotics, 1.43 mM serine and an inducer (if necessary) in 10 mL glass vials with PTFE-silicone septa screw caps. Headspace in the vials was replaced with 99% argon and 1% oxygen gas (Airgas, MA USA) using a vacuum manifold. Acetylene was injected to 10% (vol/vol) into each culture vial to begin the reaction. The acetylene reduction was carried out for 20 h at 30°C with shaking at 250 rpm, followed by quenching via the addition of 0.5 mL of 4 M NaOH to each vial.

2.4.14 Nitrogenase assays (Rhizobium strains).

Cultures were initiated by inoculating a single colony into 0.5 mL of TY medium supplemented with appropriate antibiotics in 96-deepwell plates and grown overnight at 30°C, 900 rpm in a Multitron incubator. 5 µl of overnight cultures was diluted into 500 µl of UMS medium with 30 mM succinate, 10 mM sucrose, and 10 mM NH_4Cl and appropriate antibiotics in 96-deepwell and incubated for 24 h at 30°C, 900 rpm in a Multitron incubator. Cultures were diluted to an OD_{600} of 0.4 into 2 mL of UMS medium plus 30 mM succinate and 10 mM sucrose supplemented with appropriate antibiotics, 1.43 mM serine and an inducer (if necessary) in 10 mL glass vials with PTFE-silicone septa screw caps. Headspace in the vials was replaced with 99% argon and 1% oxygen gas using a vacuum manifold. Acetylene was injected to 10% (vol/vol) into each culture vial to begin the reaction. The acetylene reduction was carried out for 20 h at 30°C with shaking at 250 rpm, followed by quenching via the addition of 0.5 mL of 4 M NaOH to each vial.

2.5 Supplemental Information

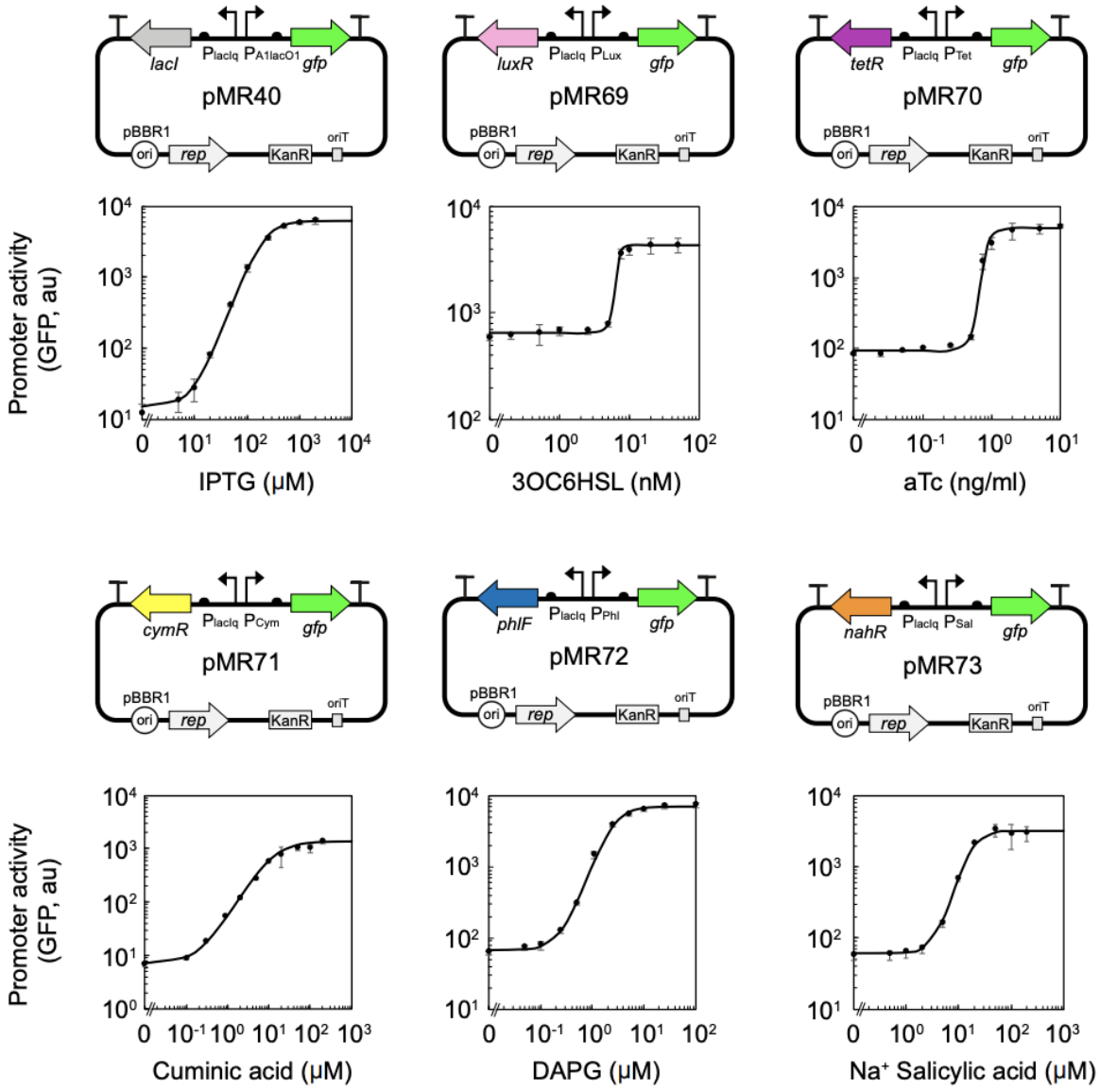

Figure 2.2 Response functions for sensors in Rhizobium sp. IRGB74
Plasmids used to characterize the sensors are shown on top of each panel. The sensors based on LacI-PA1lacO1, LuxR-PLux, TetRPTet, CymR-PCym, PhlF-PPhl, and NahR-PSal showed 398-fold, 7-fold, 62-fold, 199-fold, 120-fold, and 53-fold induction, respectively. Error bars represent standard deviation from three independent experiments on different days. Experimental details are provided in Methods.

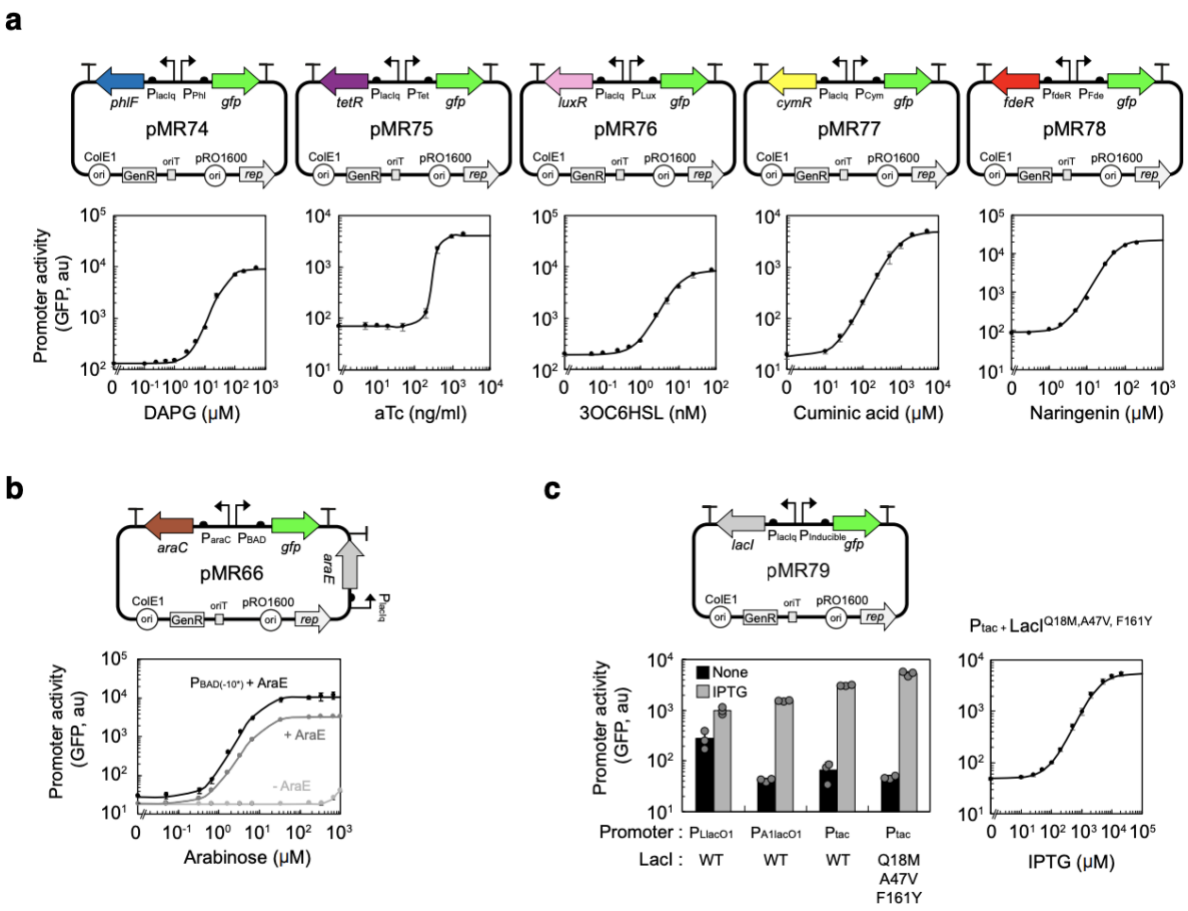

Figure 2.3 Response functions for sensors in P. protegens Pf-5
Plasmids used to characterize the sensors are shown on top of each panel. (a) Inducible promoter characterization in P. protegens Pf-5. The sensors based on PhlF-PPhl, TetR-PTet, LuxRPLux, CymR-PCym, and FdeR-PFde showed 45-fold, 61-fold, 42-fold, 544-fold, and 212-fold induction, respectively. (b) Optimization of the arabinose-inducible systems. Constitutive expression of a plasmid borne AraE transporter lowered a dissociation constant of arabinose (dark gray). A mutation in the -10 region (TACTGT to TATATT) of the PBAD promoter increased promoter strength (black). The optimized sensor showed 405-fold induction. (c) Optimization of the IPTG-inducible systems. The IPTG-inducible promoters were induced by 1 mM IPTG (left). The combination of the Ptac promoter and the LacI(Q18M/A47V/F161Y) protein yielded an expression range of 110-fold (right). Error bars represent standard deviation from three independent experiments on different days.

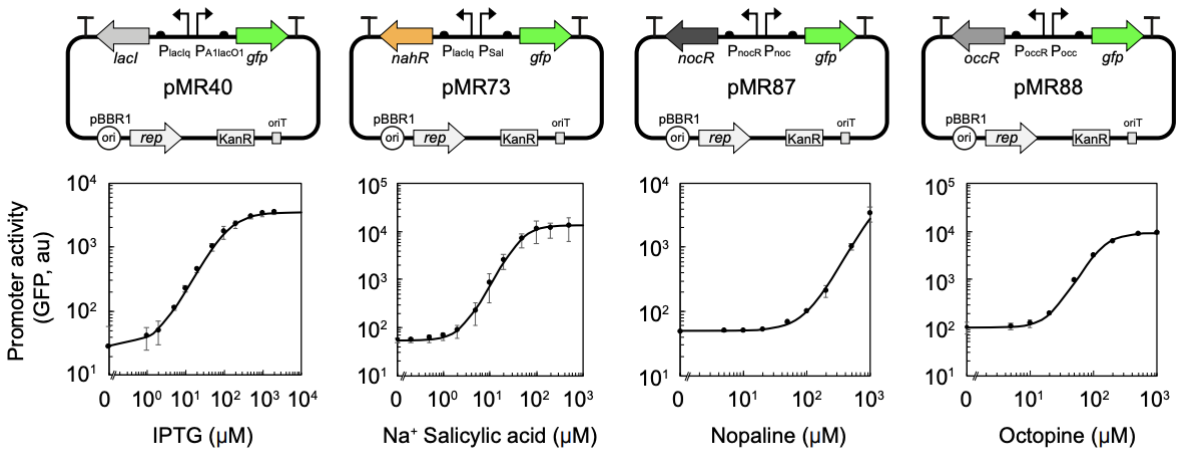

Figure 2.4 Response functions for sensors in A. caulinodans ORS571
Plasmids used to characterize the sensors are shown on top of each panel. The sensors based on LacI-PA1lacO1, NahR-PSal, NocR-Pnoc, and OccR-Pocc showed 122-fold, 238-fold, 125-fold, and 87-fold induction, respectively. Error bars represent standard deviation from three independent experiments on different days.

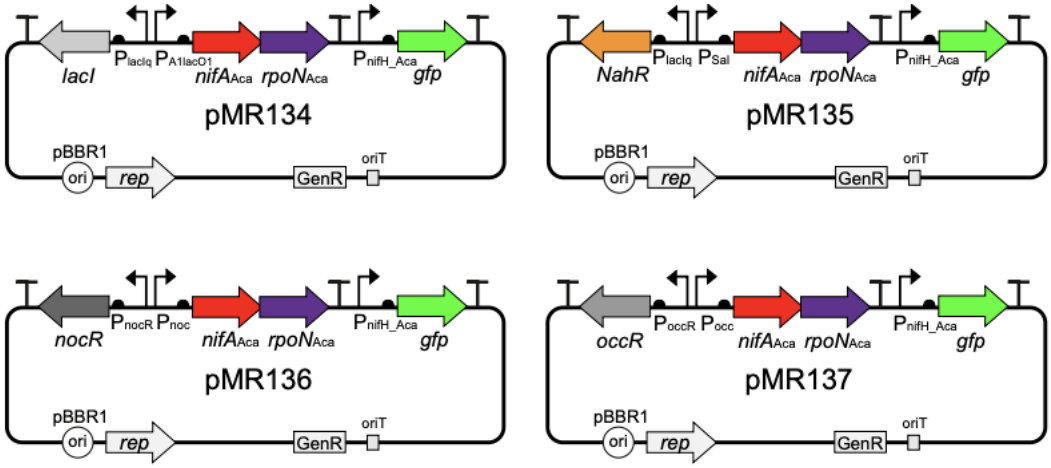

Figure 2.5 Plasmid maps for inducible expression of NifA/RpoN in A. caulinodans ORS571
Plasmid maps used to assess the effect of inducible expression of NifA/RpoN on the expression of the nifH promoter in A. caulinodans ORS571.

3 Engineering microbe-to-plant communication

This chapter is comprised of sections from a manuscript in preparation cowritten with Christopher Voigt. Qiguo Yu, Alice Boo, Alexander Pfotenhauer, Angela Belcher, Scott Lenaghan, and Neal Stewart are also authors on the paper.

3.1 Introduction

Communication between plants and microbes consists of a rich language of chemicals that transmit messages by diffusion[644-647] Plants release 100s of molecules from their roots, requiring up to 50% of the plant's photosynthetic output[648-650]. The exudate composition spans small volatile organic acids, flavonoids, lipids, oligosaccharides, peptides and proteins[646, 650-653]. Bacteria receive these signals and can respond with their own. An example of this back-and-forth chatter underlies the building of N2-fixing root nodules, where the plant releases flavonoids, the bacterium responds with nod factors, the plant grows the nodule and the concentrated bacteria respond with quorum signals sensed by both[645, 654]. The soil microbiome consists of thousands of bacteria and fungi that communicate with each other and the plant, creating a large

web of signals. Manipulating their communication has been used to improve crop yield and stress resilience; however, system complexity limits predictability[645, 652, 655-659]. The ability to access well-defined communication channels in engineering projects would facilitate the design of distributed functions across an agriculture system. For example, bacterial sentinels could survey the soil using arrays of genetically-encoded sensors and circuits and transmit the information to the plant (Figure 3.1).

Synthetic Biology projects often harness these communication signals when the spatial or temporal coordination of cells is required[660, 661]. A communication channel consists of a genetically-encoded "sender" device that produces a diffusible small molecule and "receiver" device that responds to it[662-664]. The term "device" refers to a transcriptional signal serving as the input (sender) or output (receiver), which makes it possible to connect to other devices to build a larger system. Commonly, the chemical signals are acetyl-homoserine lactones (AHLs) gleaned from quorum sensing systems in Gram-negative bacteria. The AHL is produced by a single enzyme and binds to a regulatory protein. The specificity of binding to the regulator is determined by the length of the acyl chain and different lengths have been used to build multiple non-interfering channels[665-668]. This language has been used for a plethora of projects, from stabilizing biofilm consortia to distributed computing to timing metabolic functions in a bioreactor[661, 669-678]. Eukaryotic communication channels have been developed based on different chemical signals, including peptides and pheromones[679-681].

Synthetic plant-to-microbe communication channels have been developed[682]. Plant sender devices can be built by introducing heterologous enzymes that make the roots secrete a new chemical signal[469, 607], the receiver for which is put in the bacterium. An early example created a sender in tobacco by introducing an AHL-synthesizing enzyme, which could induce root-proximal *Escherichia coli* carrying an AHL receiver[683]. Similarly, a sender in barley was built by introducing two prokaryotic genes to make scyllo-inosamine (SI), the receiver for which is the SI-binding MocB regulator, placed in in the soil bacterium *Azorhizobium caulidurans*[682, 684].

Microbe-to-plant communication is more challenging because it requires building a sensor in that plant with a low limit-of-detection. In general, constructing sensors in eukaryotes is difficult and this is compounded in plants due to even slower engineering cycles, tissue- and cell-type specific expression, fewer part libraries, chromosome context effects and complex molecular transport[11, 301, 362, 363, 400, 685]. Genetic sensors have been built that respond to ethanol, tetracycline, steroids, insecticide, trinitrotoluene (TNT), copper and fentanyl[287, 335, 685-692]. However, these are not good potential communication signals because of issues with specificity, diffusion, production by bacteria or high limits of detection. Therefore, it is a challenge to identify an appropriate chemical signal that can both be produced by bacteria and sensed by a plant.

Because of their role in establishing symbiotic relationships, AHLs have been proposed to be examples of interkingdom communication between plants and bacteria[646, 654, 693]. Soil bacteria use AHLs to communicate amongst themselves, with some able to produce up to seven AHLs and contain regulators to variants they do not make[654, 694, 695]. Biofilms on the root are abundant in AHLs, reflecting the intensity of communication that occurs there[654, 696-699]. Because of the role of AHLs in pathogenesis and symbiosis, plants have evolved means to "eavesdrop" on them to identify bacteria and respond appropriately[645, 646]. They can rapidly diffuse to the root surface from up to approximately 30 μm away, are taken up and regulate hundreds of genes via poorly-understood mechanisms[645, 700-706] The specific response depends on the plant, but short-chain AHLs change root morphology whereas long-chain (>C12) AHLs affect defense and immunity, although both can also impact energy/metabolic process, hormone production, and Ca^{2+} signaling[704, 705, 707, 708]. Plants and other microbes have the ability to interfere with AHLs by the production of degrading enzymes or chemical mimics[700, 704, 708, 709].

Rhodopseudomonas palustris is a plant-growth promoting bacterium isolated from rice paddies[710] that produces only one unique quorum signal: p-coumaric homoserine lactone (pC-HSL). In place of the aryl-group, it sources *p*-coumarate secreted from plant roots[711, 712]. No known soil bacterium has the complete pathway to pC-HSL[711], however, a synthetic pathway has been built in *E. coli* by combining *rpaI* with *rpaL*/*tal* to make endogenous p-coumarate[713]. The activator RpaR binds to pC-HSL, but not AHLs, and binds to the *rpaO* DNA[711, 713]. Some other species can make pC-HSL, but it is far less abundant than AHLs, which have been estimated to be made by 10% of the species in soil[654, 699, 711]. Evidence points to the plant response to pC-HSL being less extensive compared to AHLs, where it has only been shown to increase systemic resistance to viruses in tobacco[714, 715]. Further, there are fewer deactivating enzymes and molecular mimics in the soil[696, 716].

Here, we demonstrate programmable microbe-to-plant communication using pC-HSL sender and receiver. The sender is based on an operon containing *rpaI*/*rpaL*/*tal*, carried in the soil bacterium *Pseudomonas putida* KT2440. This strain has been proposed to be used in agriculture to promote plant growth, is non-pathogenic and does not secrete any homoserine lactones[717-719]. Because no pC-HSL receptor has been reported in plants, we built a receiver in *Arabidopsis thaliana* by constitutively expressing RpaR and building a responsive promoter. We showed that this receiver does not cross-react with AHLs. The sender can be modularly connected to different sensors and circuits and communicate the output to the plant root, demonstrated in a hydroponic system (Figure 3.1b). This work demonstrates the division of labor to move sensing to bacterial sentinels at the roots, who relay the information through a communication channel to the plant.

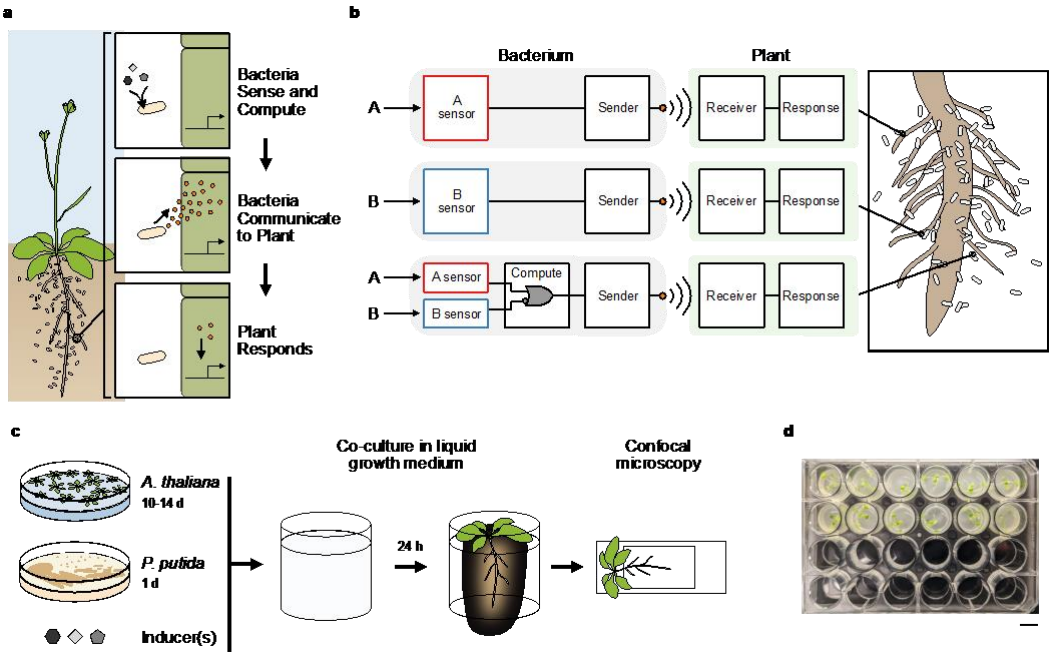

Figure 3.1 Microbe to plant communication.
a. Bacteria in the rhizosphere can be engineered to sense environmental stimuli, perform computation, and relay the signal to control a plant response. **b.** The genetic engineering of sensors, computation, senders, receivers, and response is modular and can be divided across bacteria and plant cells. **c.** Generalized scheme for hydroponic to imaging workflow. **d.** *Arabidopsis thaliana* plants grown in coculture with *Pseudomonas putida* in a 24-well plate. Scale bar represents 1 cm.

3.2 Results

3.2.1 Development of quorum receiver devices in Arabidopsis

To create a channel of microbe-to-plant communication, we first set out to create quorum receiver devices in plants. The design used quorum sensing systems from *Vibrio fischeri*, *Rhizobium leguminosarum*, *Pseudomonas aeruginosa*, and *Rhodopseudomonas palustris*, that were previously evolved to have additional chemical orthogonality (*luxR*^{AM}, *cinR*^{AM2}, *lasR*^{AM}, and *rpaR*^{AM2}, respectively) [435, 720]. For these prokaryotic transcriptional regulators to function in plants, we fused the Herpes simplex virus activation domain, VP16, to the N-terminus with a flexible glycine linker. The N-terminus was selected so as not to interfere with the DNA-binding C-terminus [721]. In addition, a nuclear localization signal (NLS) was fused to the N-terminus of the VP16 activating domain to mediate access of the chimeric sensor to the reporter construct in the nuclear genome as has been done with other synthetic transcription factors in plants (Figure 3.2b). The synthetic promoters contain 3-4 copies (3 for the P_{cin} promoter, 4 for the others) of the DNA binding operator sequence for each lux-homolog directly upstream of a minimal 35S promoter. The operator sequences used have also been evolved to have minimal promoter cross-reactivity, though

promoter cross-reactivity is less of a concern in these plants since the they only contain one luxR-homolog[720]. To visualize the receiver device output, we used a *gfp* reporter with a Tobacco Mosaic Virus Ω translational enhancer to further boost expression. With this design, in the absence of HSL the chimeric sensor protein is expressed and exists as a monomer, unable to bind to the operator sequence. In the presence of the appropriate HSL, the protein dimerizes, creating a conformational change that allows the DNA binding domain to target the operator sequence. Once bound, the viral VP16 domain recruits RNA polymerase and drives expression of GFP in the plant cell (Figure 3.2a).

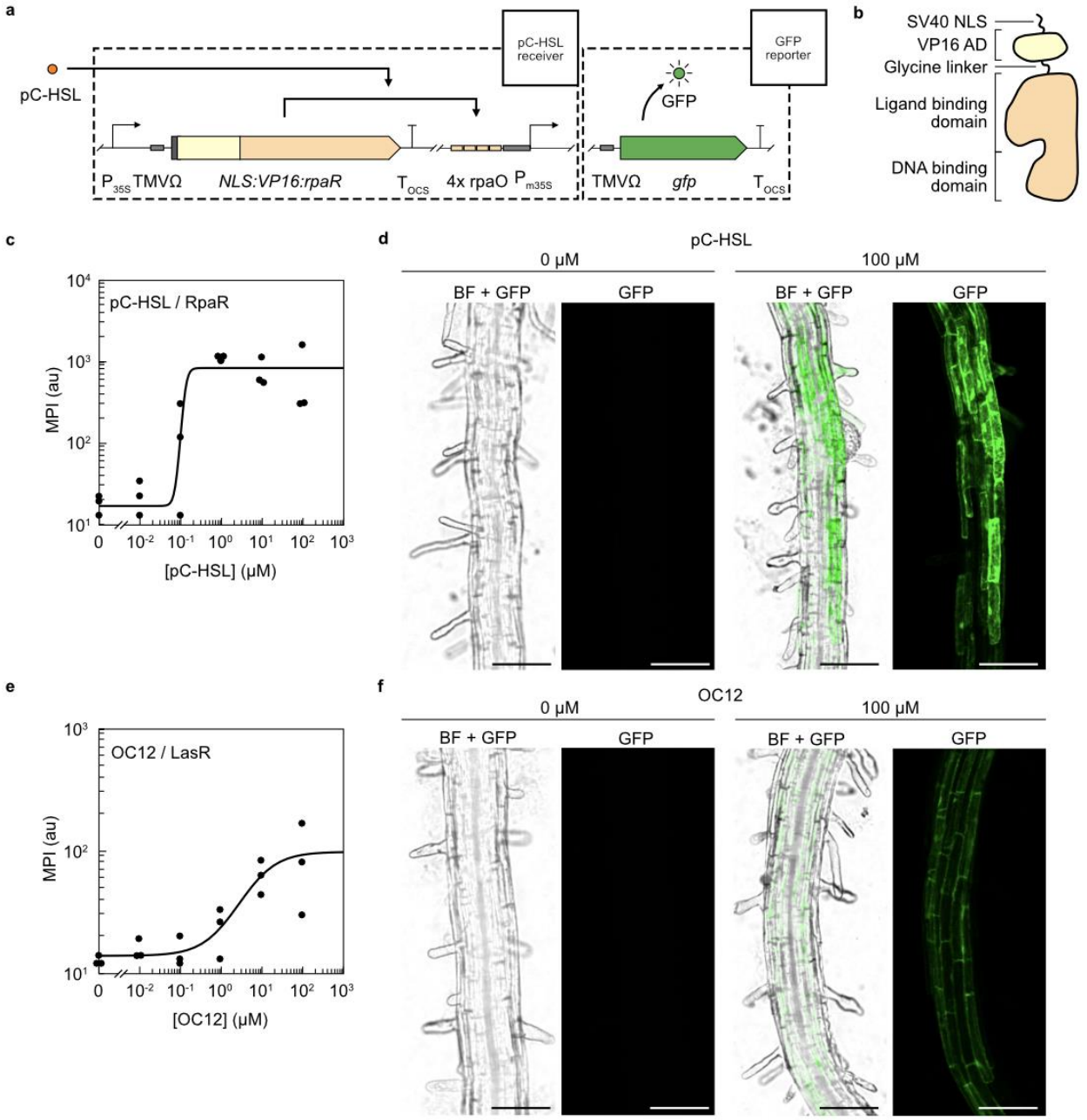

Figure 3.2 HSL receiver architecture and characterization in planta.
a. HSL receiver genetic architecture. The schematic is illustrated with the pC-HSL / *rpaR / rpaO* quorum system, yet the other HSL receiver lines had the same design. TMVΩ: Tobacco mosaic virus Ω translational enhancer, NLS: nuclear localization signal. **b.** Chimeric protein showing fusion locations. SV40: simian virus 40, AD: activation domain. **c,e.** *A. thaliana* pC-HSL receiver root GFP fluorescence MPI after 24 hours of induction with different concentrations of HSL. Data points represent experiments repeated on different days with different plants (n=3). The response profile is fit to a Hill equation using Excel. **d,f.** Fluorescent and composite images of *A. thaliana* HSL receiver roots. Images are representative of experiments performed on 3 different days with different plants. Brightness adjusted for clarity (min, max): GFP (150, 4095). Scale bars represent 100 μm.

Quorum receiver *Arabidopsis thaliana* plants were created using the *Agrobacterium tumefaciens* floral dip method. Multiple *A. thaliana* lines homozygous for each of the receivers were created and all were phenotypically indistinguishable from wildtype plants. We next screened multiple independent lines for each quorum receiver to determine the top performing lines, which we defined as having the greatest dynamic range of fluorescent output between uninduced and induced states. Plants were germinated and grown on agar plates for 10 days in a growth chamber before being transiently grown in a hydroponic system. A hydroponic system allowed for us to easily adjust the media composition and expose the plant roots more uniformly to inducer than agar plates would have allowed. Additionally, we used a 24-well plate with one plant per well to increase the throughput of conditions in which the plants were exposed. With this system in place, we then tested each HSL receiver line against each of the four HSL inducers (pC-HSL, OC12, OC6, and OHC14) at a maximum concentration of 100 μM for 24 hours.

We quantified the GFP reporter output of the HSL receiver lines using fluorescence images of the root tissue. A custom ImageJ macro was used to quantify the mean pixel intensity (MPI) of the GFP channel across the root tissue sections (Figure 3.7). To enable comparisons between different days and plant lines, we used the same laser intensities and microscope settings for all experiments. Using this approach, we saw a 25 and 6-fold increase in MPI of the GFP channel between uninduced and induced plants for the top performing pC-HSL receiver and OC12 receiver lines, respectively. None of the non-cognate HSLs resulted in a change in MPI for these two lines suggesting that the receiver lines remain chemically orthogonal (Figure 3.2e). Additionally, when uninduced, the top performing lines had a MPI similar to wildtype controls suggesting little to no background expression of the GFP reporter (Figure 3.9). The top performing OHC14 receiver line showed a 2-fold change in MPI between uninduced and fully induced states. None of the OC6 receiver lines showed significant fold change in MPI and were therefore not tested further.

To determine the detection limit of the top pC-HSL receiver and OC12 receiver lines, they were induced with 10-fold increments of pC-HSL and OC12, respectively. The top performing pC-HSL receiver line was fully induced (>25 fold change in MPI versus the uninduced state) with 1 μM or greater of pC-HSL (Figure 3.2f). The fluorescent signal did not appear uniformly distributed across the root tissue (Figure 3.2g). The top OC12 receiver line was less sensitive, requiring 100 μM of OC12 for max GFP expression and a 7.3-fold change in MPI compared to the uninduced plant (Figure 3.2h-i). While still detectable, the relatively low fluorescence at even the highest induction levels tested of the OC12 receiver and OHC14 receiver lines made them poor candidates for microbe-to-plant communication and were thus not used in future experiments. In contrast, the low threshold of detection and the high fluorescent output of the pC-HSL receiver lines made them a prime candidate for creating a microbe-to-plant signaling channel.

3.2.2 A Pseudomonas putida sender device

The soil bacterium *Pseudomonas putida* was engineered to produce pC-HSL under different regulatory control, creating a pC-HSL sender device. *P. putida* was selected as the chassis due to its abundance in the rhizosphere, broad metabolic capabilities, and amenability to genetic engineering. The biosynthetic pathway for pC-HSL starting from a common metabolite L-tyrosine can be completed with the addition of three genes - *tal*, a tyrosine ammonia lyase which converts L-tyrosine to *p*-coumarate; rpaL, a 4-coumarate-CoA ligase which converts *p*-coumarate to *p*-coumaroyl-CoA; and *rpaI*, the acyl homoserine synthase which converts *p*-coumaroyl-CoA into pC-HSL. To test *P. putida* as a sender device, the pC-HSL biosynthesis operon (*rpaI rpaL tal*) was expressed constitutively from a strong promoter on a relatively high copy number plasmid. In addition, an *mCherry* reporter was also included in the operon to serve as a visual proxy for pC-HSL production (Figure 3.3a).

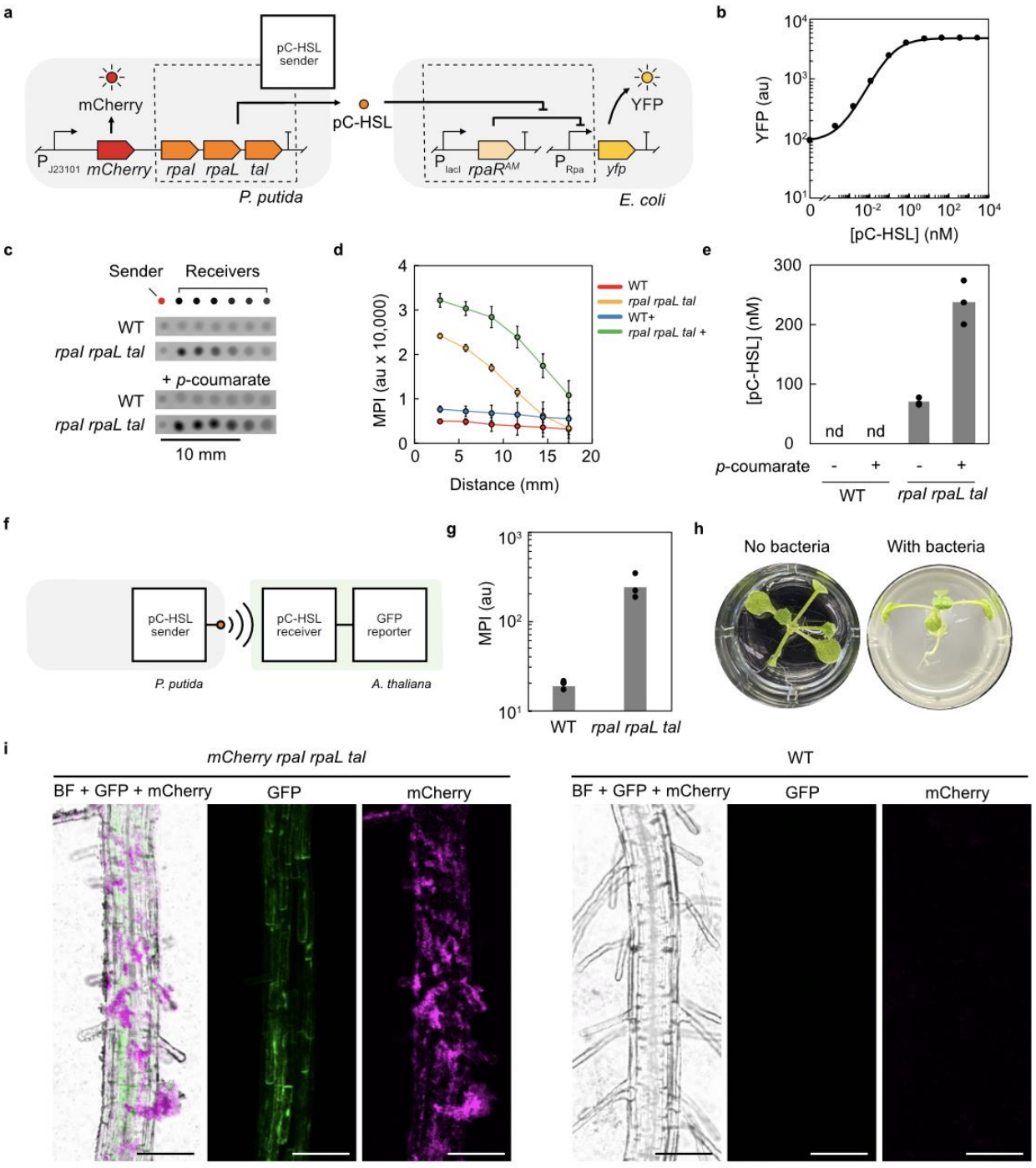

Figure 3.3 Biosynthesis of pC-HSL for bacteria-to-plant communication.
a. Schematic showing *P. putida* pC-HSL sender device architecture and *E. coli* pC-HSL receiver. **b.** Dose-response curve of *E. coli* pC-HSL receiver with a YFP output. The median YFP fluorescence of at least 10,000 cells is plotted. Error bars represent the standard deviation from three replicates. The response profile is fit to a Hill equation using Excel. **c.** Colonies of *P. putida* pC-HSL sender or WT and *E. coli* pC-HSL receiver printed on agar plates and grown for 24 hours. Images captured using ChemDoc and darker color indicates higher YFP expression. *p*-coumarate added at 100 µM where denoted. **d.** Quantification of YFP fluorescence of colonies in 3c. Mean pixel intensity calculated using a circle with a 13-pixel diameter centered on each colony using ImageJ. + denotes conditions with the addition of 100 µM *p*-coumarate. **e.**

Supernatant dilutions of *P. putida* pC-HSL sender used to induce *E. coli* pC-HSL receiver and estimate pC-HSL production using the dose response in 3b. Production was increased with the addition of 100 µM *p*-coumarate. Data points represent estimated production from 3 replicates. nd: not detectable. **f.** Schematic representation of *P. putida* pC-HSL sender and *A. thaliana* pC-HSL receiver grown in coculture. **g.** *A. thaliana* pC-HSL receiver root GFP fluorescence MPI after 24 hours cocultured with WT or pC-HSL sender *P. putida*. 100 µM *p*-coumarate added to growth media. Data points represent experiments repeated on different days with different plants (n=3). **h.** *A. thaliana* pC-HSL receiver plants grown in hydroponic coculture media for 24 hours. Shown with and without *P. putida* inoculated at a starting OD of 0.01. **i.** Fluorescent and composite images of *A. thaliana* pC-HSL receiver roots cocultured as depicted in 3f. Images are representative of experiments performed on different days with different plants. Brightness adjusted for clarity (min, max): GFP (150, 4095); mCherry (120, 1000). Scale bars represent 100 µm.

The production of pC-HSL by *P. putida* was confirmed in two ways, both using a coculture with *E. coli* receiver cells containing a $rpaR^{AM2}$ sensor driving expression of *yfp*. *P. putida* sender cells were printed near multiple *E. coli* receiver cells on agar plants and grown spatially separated. After 24 hours of growth, the plates were imaged using a gel imager to capture the fluorescence signal. The YFP fluorescence of the *E coli* receiver cells was strongest closest to the *P. putida* sender cells and decreased with distance, indicating that the receives were being induced by pC-HSL from the sender (Figure 3.3c). We quantified a 4.1-fold increase in YFP fluorescence in the nearest receiver cells as measured by mean pixel intensity of the colony compared to WT (Figure 3.3d). These data suggest that pC-HSL was produced by the sender cells and that it could diffuse freely through the agar plate. Wildtype *P. putida* did not cause any significant change in MPI of any receiver cells (Figure 3.3c-d). Addition of *p*-coumarate to the agar plate resulted in an increase in YFP signal from all the *E. coli* receiver colonies grown near *P. putida* sender cells, suggesting that more pC-HSL was being produced and it was therefore diffusing farther from the sender cell. In addition, these results show that the endogenous precursors are limiting in the production of pC-HSL by *P. putida* sender cells and that supplementing the media with *p*-coumarate is an effective way to boost production.

In addition to validating pC-HSL production on agar plates, dilutions of the supernatant from *P. putida* pC-HSL sender cells grown in liquid culture for 24 hours were used to induce liquid cultures of the *E. coli* sensor cells. By using dilutions of the supernatant and the known dose response of the *E. coli* sensor cells (Figure 3.3b), we could estimate that the production of pC-HSL was equivalent to inducing the sensor cells with 68 nM of pC-HSL over the 24 hour growth period (Figure 3.3d). In agreement with the results from the agar assay, we found that the addition of 100 µM *p*-coumarate to the growth media increased the production of pC-HSL 3.5-fold to an equivalent of inducing the sensor cells with 235 nM over the 24 hour growth period. Collectively, these results show that the *P. putida* constitutive sender device is capable of producing pC-HSL and that the production is boosted with the addition of *p*-coumarate precursor.

3.2.3 Bacteria-to-plant signal relay

We next explored whether the *P. putida* pC-HSL sender and *A. thaliana* pC-HSL receiver devices could be connected to create a bacteria-to-plant signal relay (Figure 3.3f). We first adjusted the media composition of the hydroponic system to allow for the growth of both *P. putida* and *A. thaliana* in coculture (Figure 3.6). There was a slight change in phenotype of plants grown in coculture with bacteria after 24 hours, usually in the form of modest hyperhydricity though we don't believe that it affected the fluorescence in the root tissue (Figure 3.3h). We then grew the top pC-HSL receiver plant line with the constitutive *P. putida* pC-HSL sender cells or wildtype *P. putida*. After imaging the roots, we calculated a 12.7-fold increase in MPI of GFP fluorescence of the pC-HSL receiver roots grown with the pC-HSL sender compared to coculture with wildtype *P. putida* (Figure 3.3g). These data support the ability of *P. putida* pC-HSL senders to induce expression of the GFP reporter in the pC-HSL receiver *A. thaliana* lines. The mCherry reporter of the sender cells could also be used to visualize *P. putida* surrounding the root tissue (Figure 3.3i). To our knowledge, this is the first demonstration of engineered bacteria-to-plant communication.

We next sought out to use bacteria grown in coculture to relay the detection of an external stimuli to a plant. First, *cymR*, *lacI*, *tetR*, and *phlF* transcriptional regulators were ported to and characterized in *P. putida* to create a cumate, IPTG, aTc, and DAPG sensor, respectively[435] (Figure 3.4a).

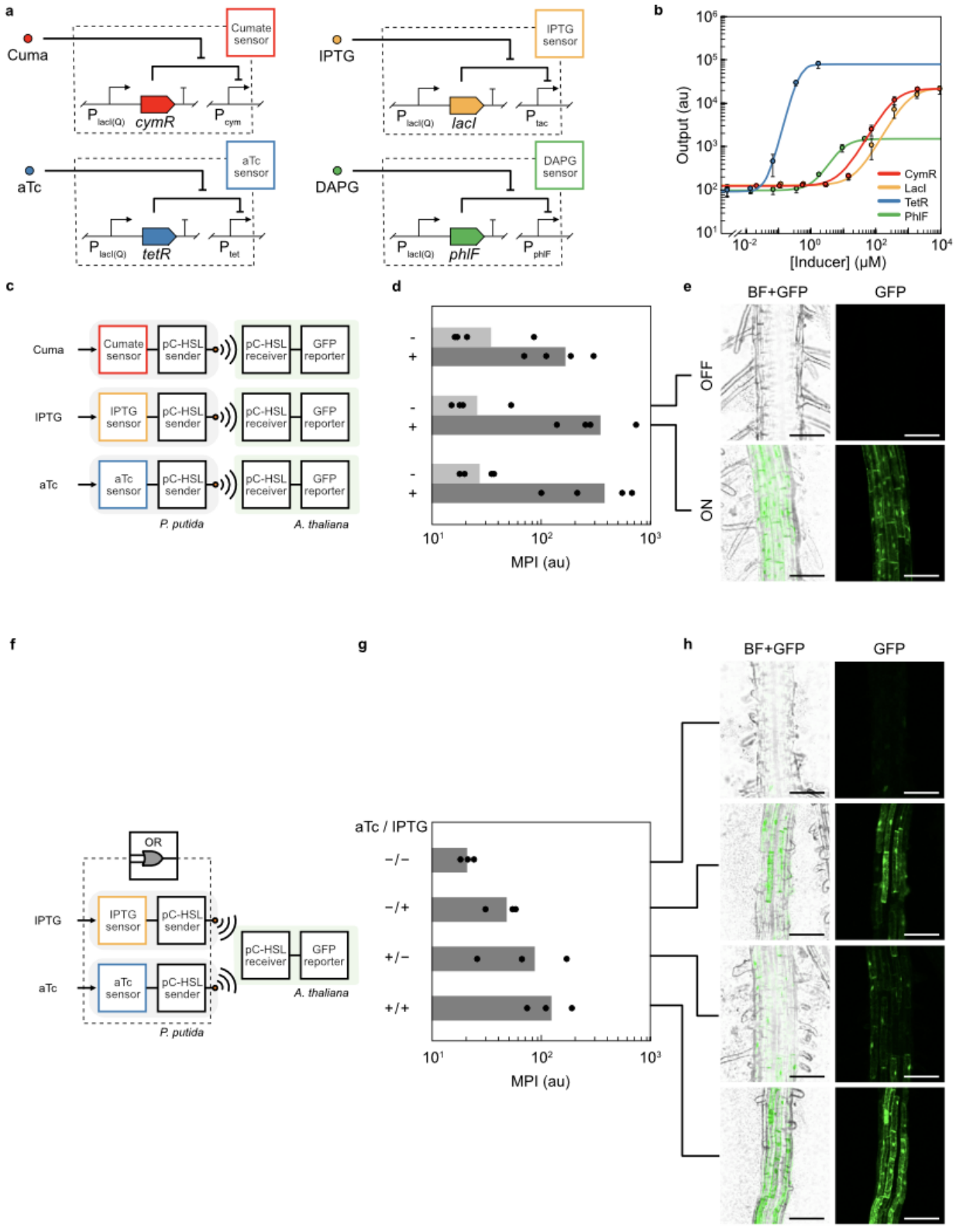

Figure 3.4 Relaying detection of external stimuli to a plant with bacteria sensors.
a. Schematics showing the sensor architecture for four different small molecule sensors. **b.** Dose-response curves of *P. putida* sensors with a YFP output. The median YFP fluorescence of at least 10,000 cells is

plotted. Error bars represent the standard deviation from replicates on different days (n=3). The response profile is fit to a Hill equation using Excel. Histogram distributions provided in Figure 3.12. **c.** Schematic representation of *P. putida* sensor-sender cells and *A. thaliana* pC-HSL receiver plants grown in coculture. **d.** *A. thaliana* pC-HSL receiver root GFP fluorescence MPI after 24 hours cocultured with *P. putida* sensor-sender cells as depicted in 4c. Inputs correspond to the absence (-) or presence (+) of: 500 µM cumate (Cuma), 2 mM IPTG, 1 µM aTc, and 50 µM DAPG. Data points represent experiments repeated on different days with different plants (n=4). 100 µM *p*-coumarate added to growth media. **e.** Fluorescent and composite images of *A. thaliana* pC-HSL receiver roots cocultured as depicted in 4c. Images are representative of experiments performed on different days with different plants. Brightness adjusted for clarity (min, max): GFP (150, 4095). Scale bars represent 100 µm. **f.** Schematic representation of two different *P. putida* sensor-sender cells and *A. thaliana* pC-HSL receiver grown in coculture to create OR gate logic. **g.** *A. thaliana* pC-HSL receiver root GFP fluorescence MPI after 24 hours cocultured with two *P. putida* sensor-sender cells as depicted in 4f. Inputs correspond to the absence (-) or presence (+) of 2 mM IPTG and 1 µM aTc. Data points represent experiments repeated on different days with different plants (n=3). 100 µM *p*-coumarate added to growth media. **h.** Fluorescent and composite images of *A. thaliana* pC-HSL receiver roots cocultured as depicted in 4f. Images are representative of experiments performed on different days with different plants. Brightness adjusted for clarity (min, max): GFP (150, 4095). Scale bars represent 100 µm.

We tested the response of each of these sensors using a YFP reporter. Sensor cells were grown in LB with appropriate inducers after which the YFP fluorescence was measured with flow cytometry. The dynamic ranges of the sensors ranged from 10-fold for *phlF* (DAPG sensor) to 735-fold for *tetR* (aTc sensor) (Figure 3.4b, Figure 3.12). *P. putida* sensor-sender strains were then constructed using each of the sensors to control the pC-HSL sender device (Figure 3.4c). The *P. putida* sensor-sender cells were grown in coculture with the pC-HSL receiving plant line and the appropriate small molecule inducer at the optimal concentration as determined by the dose response for each sensor: 500 µM cumate (Cuma), 2 mM IPTG, 1 µM aTc, and 50 µM DAPG. After 24 hours of coculture, the pC-HSL receiver plant roots were imaged and the GFP reporter expression was quantified as MPI. Each of the *P. putida* sensor-sender strains was able to relay detection of its cognate small molecule to the plant as indicated by an increase in root MPI when induced (Figure 3.4d). The cumate sensor resulted in the smallest fold change in MPI at 4.8 while the aTc sensor had the greatest at 13.9. Figure 3.4e shows representative images of IPTG being detected by the *P. putida* IPTG sensor-sender, controlling the production of the pC-HSL bacteria-to-plant relay, and the plant GFP reporter being expressed in response. These results show the pC-HSL relay can be used to create plants that can effectively sense any small molecule that a bacteria can be engineered to detect.

3.2.4 Performing computation using an engineered bacterial consortium

Next, we employed a consortium of engineered bacteria grown in co-culture as an alternative approach to engineering a system with OR gate logic (Figure 3.4f). In other words, rather than engineering an OR gate directly into a plant or bacteria, the presence of two sender-sensor bacteria simultaneously means that the presence of either inducer should result in reporter expression, hence replicating OR gate

logic. Here, *P. putida* sensor-sender strains capable of relaying detection of IPTG and aTc were grown with the pC-HSL receiver plant line with neither, one, or both inducers present. The MPI of the receiver plant remained at baseline levels when neither IPTG or aTc were added (Figure 3.4g-h). Addition of 1 µM aTc, 2mM IPTG, or both 1µM aTc and 2 mM IPTG resulted in a 2.3-, 4.1-, and 5.9-fold increase in MPI, respectively. These data show that the response of the plant to the different combinations of inducers matched the desired OR gate response, without having to directly engineer an OR gate. It was evident that the MPI was lower for the consortium ON states than it was when using individual sensor lines. We believe this is due to the starting inoculum containing half the number of each sensor cell since the total starting number of bacterial cells was constant for all coculture experiments, regardless of if one or two strains were inoculated. The use of an engineered consortia show that genetic engineering approaches can be split not only between plants and bacteria, but further divided across multiple strains of bacteria to create desired holistic responses.

3.3 Discussion

This work demonstrates a programmable mode of communication from a bacterium to a plant. The communication signal is defined as pC-HSL, a sender of which is built in the bacterium and a receiver in the plant. Characterizing them as transcriptional devices simplifies their incorporation into larger genetic designs. Here, we used simple sensors in the bacterium (aTc, IPTG, cumate), but the sender could be connected to any sensor whose promoter output has a matching dynamic range, including those that respond to toxins, pollutants, nutrients, pathogens, agrochemicals and any other stimuli relevant to agriculture[346, 663, 722, 723]. The sender could also be connected to the output promoter of a genetic circuit that integrates information from multiple sensors or implements a dynamic response[242]. Similarly, the pC-HSL receiver was used here to control reporter expression, but it could be connected to metabolic pathways or transcription factors to control morphology[724]. This enables sense-and-respond systems to be distributed, where a bacterial sentinel receives information (e.g., toxin), transmits it to the plant, which then turns on a response (e.g., detoxification pathway). This allows complex sensing and circuitry to be moved to the bacterium, where it is easier to build[242, 435], and it reduces the number of genes that need to be stacked in the plant, thus reducing the growth burden and reducing regulatory concerns. Further, when combined with the plant-to-microbe communication developed previously[725], this results in the ability to program two-way communication to coordinate interkingdom cooperation, such as the establishment of synthetic symbiosis[483, 583].

The plant pC-HSL receiver operates in a response regime consistent with signals in the rhizome at the root. The detection limit is 100 nM, which is similar to the 20-50 nM detection limit of native *R.*

palustris, which produces up 1-10 µM pC-HSL in culture[711]. This threshold is higher than nod factor (0.01 – 1 nM) perception by legumes[645], but at the low end of the 1 nM to 100 µM AHLs required to stimulate root growth or promote protective functions in various plant species[703, 704, 708, 726]. It is also much lower than the 5 µM pC-HSL used to induce defense genes in tobacco[714]. The pC-HSL productivity by engineered *P. putida* KT2440 in culture can be estimated as ~250nM over 24 hours which meets the detection limit within the induction period.

The selection of pC-HSL came as the result of a process of elimination, as opposed to the top-down identification of the perfect communication molecule. The constraint was that the molecule had to both be producible by a bacterium and sensed by a plant. Transport is also an important consideration, as the molecule has to be exported through the prokaryotic membrane and taken up by the root, ultimately entering the nucleus. Prior to selecting pC-HSL, we tested other candidates (not shown). Lipo-chitooligosaccharides (nod factors) were considered, but we had difficulty both in producing high titers in bacteria and creating a sensor that could be moved to non-legumes. There is an indigo sensor for plants [727] and we could make it at high titer in bacteria using published pathways [728], but we could not export it without lysis. Another mode of interkingdom signaling is the production of plant hormones, such as auxins, by growth-promoting bacteria[729]. However, these have extensive effects on plant growth and gene expression, and are ubiquitous across plant species. Using the same plant promoter scaffold we used in this work, we also tried to make sensors for DAPG (PhlF), 3OC6-HSL (LuxR), 3OC8-HSL (TraR) and 3OHC14-HSL (CinR), all of which could be produced by bacteria, but they failed. The only AHL that we could make a sensor for was 3OC12-HSL (LasR), which could offer a second orthogonal communication channel, noting that its response has a lower dynamic range (Figure 3.2e) and it is known to induce Arabidopsis genes[704, 708, 730].

Rare homoserine lactones may be the ideal molecules for interkingdom signaling. Homoserine lactones can transport across the walls of many cell types, can be produced at sufficient titer with a few enzymes, are non-toxic, and bind to well-defined regulators. AHLs may be too common to enable specificity between an engineered plant and microbe and have too systematic of impact on plant gene expression, but there are many other structures from which to choose. This includes unusual branched-chain groups, as well as other aroyl- groups (*e.g.*, cinnomoyl-HSL), as well as mimicking compounds[625, 653, 731] and there are likely many more that have yet to be discovered[654, 732]. Non-natural HSLs have been built with synthetic organic chemistry, including sulfonyl, aroyl and alkanoyl-HSLs[715, 733] and biosynthetic retrosynthesis tools[734, 735] could be applied to build senders and directed evolution of LuxR homologues to make receivers[736-738]. Homoserine lactones have the possibility of being inactivated by lactases in the soil[716], but it is noteworthy that there are plant-based systems that react to the version of these molecules with open rings, whereas bacterial regulators require the rings to be closed[693]. These molecules open up the possibility for more soil stability and less interference from other members of the microbiota. Similarly, all natural

homoserine lactones are L-isomers and D-isomers are not biologically active in plants[739, 740]. Synthetic pathways to the D-isomers could reduce crosstalk with plant signaling, but this would require engineering new regulators.

Microbes beneficial to crops have long been used in agriculture. Typically, one species has been used or few have been combined into an artificial consortium. Building such consortia has been ad hoc, and it has been noted that the positive effects of multiple species often do not combine additively[697]. This work lays the foundation for using bacteria grown in proximity to a plant to perform tasks such as monitoring soil nutrient content, sensing pathogens, or detecting environmental contaminants. For example, nitrogen and phosphorous sensors in bacteria can be used to monitor the availability of nutrients in the soil, relay the amounts to the plant, and the plant could report high or low quantities to farmers. Alternatively, bacteria can be engineered to detect the presence of soil pathogens, such as *Botrytis cinerea* (causative agent of grey mold or botrytis bunch rot common in grapes) or *Gaeumannomyces graminis* var. *tritici* (causative agent of take-all in wheat, barley and other grasses), that could decimate yields, relay their presence to a plant to turn on defense genes (Induced Systemic Response) or to produce targeted antimicrobials to fend off the attack, such as 2,4-diacetylphloroglucinol[741, 742]. In addition to agricultural applications, this approach could be applied to build plant sentinels. Moving the sensing to root-associated bacteria allows the same plant to be used to detect different signals simply by swapping the engineered bacterium with which it is partnered that contains the new sensor. These microbially mediated phytosensors can perform incognito detection of chemical or biological warfare agents in the soil[743-745].

These tasks would be either too complex to engineer directly into plants or could be distributed to soil microbes to reduce growth impacts on the plants and ease regulatory concerns. By viewing the plant-microbe community holistically, we can select the organism best suited for a particular task and use orthogonal channels of communication to coordinate the population to accomplish complex engineering tasks.

3.4 Methods

3.4.1 Bacterial strains and growth conditions.

Escherichia coli NEB 10ß (New England BioLabs C3019I) were used to clone all plasmids. *Pseudomonas putida* KT2440 (ATCC 47054) were used for coculture experiments. *Agrobacterium tumefaciens* GV3101 (Gold Bio GV3101 Electrocompetent) were used for leaf transient assays and the floral dip method. Cells were routinely grown in LB Miller broth (Difco 244620) at 37°C for *E. coli* and 30°C for *P. putida* and *A. tumefaciens*. Antibiotics were used to maintain plasmids during routine growth: kanamycin (GoldBio K-120-10) - 35 µg/mL for *E. coli*, 50 µg/mL for *P. putida* and *A. tumefaciens;*

tetracycline – 10 µg/mL for *E. coli*, 25 µg/mL for *P. putida*; gentamycin – 15 µg/mL for *E. coli*, 50 µg/mL for *P. putida*; chloramphenicol (Alfa Aesar B20841) – 25 µg/mL for *P. putida;* rifampicin – 50 µg/mL for *A. tumefaciens*. DNA oligos and genes were ordered from Integrated DNA Technologies. Parts and plasmid maps for all constructs are provided in Supplementary Notes. Plasmids were assembled using Gibson assembly. Assembled plasmids were purified from *E. coli* NEB10ß using QIAGEN miniprep protocol and sequence verified. Plasmids were either electroporated into competent *P. putida* or *A. tumefaciens* following previously published methods[725] or conjugated using triparental mating [471].

3.4.2 Chemicals.

Ten mM stocks of HSLs were solubilized in DMF and stored at -20°C: OHC14 (3OHC14-HSL, Sigma 51481); OC6 (3OC6-HSL, Sigma K3007); OC12 (3OC12-HSL, Sigma O9139); pC-HSL (*p*-coumaroyl-HSL, Sigma 07077). Bacterial cells were induced using the following chemical stocks: DAPG (2,4-diacetylphloroglucinol, Santa Cruz sc-206518) in DMF; Cuma (cuminic acid, Millipore Sigma 268402) in ethanol; IPTG (isopropyl-ß-D-1-thiogalactopyranoside, Gold Biotechnology I2481) in water; aTc (anhydrotetracycline, Millipore Sigma 37919) in 50% (v/v) ethanol; Ara (L-arabinose, Sigma A3256) in water.

3.4.3 Sensor and circuit characterization in P. putida.

Overnight cultures were grown in 1mL LB with antibiotics in 96 deep-well plates. 0.5 µL of overnight culture was diluted into 150 µL LB with antibiotics and inducer in 96 well V-bottom plates and grown at 30°C at 1000rpm. The cells were diluted into 1x PBS (OmniPur 10x PBS liquid concentration, Millipore Sigma 6505-OP) before running flow cytometry. Fluorescence characterization was performed using a BD LSR Fortessa flow cytometer with HTS attachment (BD Biosciences). At least 10,000 events were captured for each sample. Measurements were made using a green laser (488 nm) voltage of 450 V, an FSC voltage of 750 V, and SSC voltage of 300 V. FlowJo and Cytoflow were used for analysis and gating. The median fluorescence is used to summarize cytometry distributions. Response functions were fit to a hill equation using Microsoft Excel Solver.

3.4.4 Arabidopsis thaliana Col-0.

Arabidopsis thaliana Col-0 seeds (NACS CS70000) were acquired from the Arabidopsis Biological Resource Center. When working with a small number of seeds, they were surface sterilized with 70% ethanol for 1-2min, followed by 10% bleach for 10 min, and then rinsed 5 times with water. Larger numbers of seeds were sterilized using the chlorine gas method[746]. Sterilized seeds were sown on half strength Murashige and Skoog (MS; Sigma M5519) media, adjusted to a pH of 5.7 with KOH. Plates were

made with the addition of 0.8% phytoagar (Sigma A7921). For selections, phosphinothricin (PPT; Gold Bio P-165-250) was added at a final concentration of 50 µM. Plates were sealed with micropore tape (3M 1530-0) to allow for gas exchange. Seeds were stratified at 4°C in the dark for at least 3 days before moving to a growth chamber where they were grown at 27°C in 16/8 hr light/dark cycles with a light intensity of ~100 µmol/m^2.

3.4.5 Agrobacterium floral dip.

Transgenic *Arabidopsis* lines were generated by *Agrobacterium tumefaciens*-mediated floral dip. Briefly, *A. tumefaciens* strains containing plasmids of interest were cultured in 2 mL LB containing appropriate antibiotics at 30°C and 250 r.p.m. for 2 days. This culture was used to inoculate 500 mL of LB with appropriate antibiotics and cultured an additional 24 hrs. Grown cultures were moved to 1 L bottles and pelleted by centrifugation at 4,000g for 10 min at 4°C. Pellets were resuspended by pipetting with a serological pipette in 5% (w/v) sucrose solution plus 0.02% (v/v) Silwet L-77 (PhytoTech Labs S7777). Arabidopsis inflorescences were submerged in the bacterial resuspension for 1 min with gentle agitation, removed and drip dried, and covered gently in plastic wrap before being transferred to the dark overnight. The next day transformed plants were returned to the greenhouse until they produced seeds.

Transgenic T1 seedlings were sown on moistened soil and covered with clear plastic lid until cotyledons were visible at which point they were sprayed with PPT herbicide twice per week for ~3 weeks until only resistant lines remained. T1 plants were grown to seed. T2 seeds were sown on agar plates containing 50 µM PPT. Seedlings were transferred to soil at ~3 weeks and grown to seed. Homozygous T3 lines were validated by segregation on selection plates.

3.4.6 Chemical HSL induction of HSL receiver Arabidopsis.

Stable T3 seed was surface sterilized as described above and sown on MS/2 agar plates with 1% sucrose. The plates were placed in the dark at 4°C for a 3-day striation period before being moved to a growth chamber and grown for 10–12-days. In a tissue hood, individual wells of 24-well plates (Falcon 353047) were filled with 1mL of MS/2 medium with 1% sucrose. Chemical HSL was added to appropriate wells. Finally, plants were carefully lifted from the agar plates with forceps and moved to individual wells such that the roots were entirely submerged. Each plate was covered with a lid, sealed with micropore tape, and returned to the growth chamber. After 24 hours, plates were taking from the growth chamber for imaging.

3.4.7 Bacterial induction of HSL receiver Arabidopsis

Stable T3 seed was surface sterilized as described above and sown on MS/2 agar plates with 1% sucrose. The plates were placed in the dark at 4°C for a 3-day striation period before being moved to a growth chamber under the conditions described above and grown for 10–12-days. Two days before plant inoculation, *P. putida* strains were struck out from glycerol stocks onto LB agar plates with appropriate antibiotics and grown at 30°C overnight. The next day, individual colonies were selected and inoculated into 1mL of LB media with appropriate antibiotics in 96 deep-well plates. These plates were grown overnight at 30°C at 900rpm (Infors Multitron). The day of plant inoculation, the OD_{600} was measured using a spectrometer. In a tissue hood, individual wells of 24-well plates (Falcon 353047) were filled with 750 µL of MS/2 medium with 1% sucrose and 250 µL of LB. *P. putida* was inoculated in appropriate wells to a starting OD_{600} of 0.01. For *P. putida* containing inducible control of HSL production from individual sensors, appropriate chemical inducers were added at the following concentrations: DAPG – 50 µM; Cuma – 500 µM; IPTG – 2 mM; aTc – 1 µM. For *P. putida* containing inducible control of HSL production from a circuit, appropriate chemical inducers were added at the following concentrations: aTc – 1 µM; IPTG – 2 mM. When noted, 100 µM of *p*-coumarate was also added. Finally, plants were carefully lifted from the agar plates with forceps and moved to individual wells such that the roots were entirely submerged. Each plate was covered with a lid, sealed with micropore tape, and returned to the growth chamber. After 24 hours, plants were removed from each well and placed on a microscope slide (VWR Micro Slides 48300-026) under a 22x40mm No 1 cover slip (VWR Cover glass 48393-048) for imaging.

3.4.8 Confocal microscopy

Microscopy experiments were performed using a Nikon A1R Ultra-Fast Spectral Scanning Confocal Microscope. GFP signal was captured with 488 nm excitation and xx nm emission filter. The 487.3 nm laser was held used at 10% power, HV 100, and 0 offset. Upon loading each sample, the entire root system of each plant was inspected, and images were captured only of the brightest portion. A 4x objective was used unless otherwise noted.

3.4.9 Image analysis and quantification

All images were analyzed with FIJI package of ImageJ. Bright field and fluorescent images were aligned using the Landmark Correspondences plugin. For composite images, the LUT of the fluorescent channels were inverted (Figure 3.8). For display purposes and to facilitate comparisons, the maximum and minimum brightness were adjusted in some images and the values are listed in the figure captions. Only the unaltered images were used for MPI quantification. A custom macro was used for fluorescence quantification (Figure 3.7). Briefly, a Gaussian filter is used for noise reduction followed by thresholding

at a constant value to separate root tissue from background. The mean pixel intensity of the root tissue minus the mean pixel intensity of the background is then calculated.

3.5 Supplemental Information

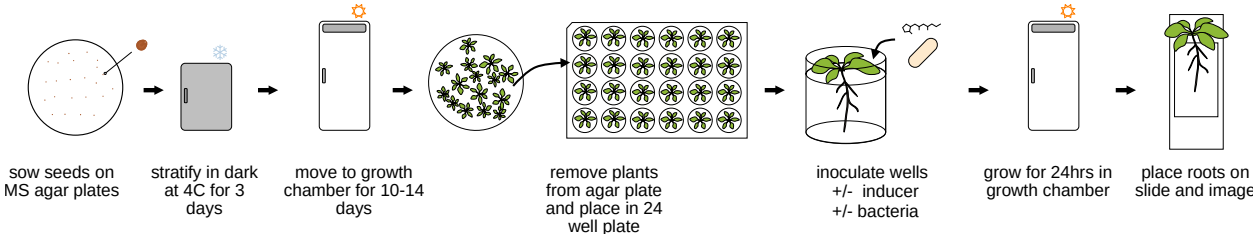

Figure 3.5 Transient hydroponic growth overview.

Seeds are surface sterilized with 70% ethanol for 1-2 min, followed by 10% bleach for 10 min, and then rinsed 5 times with water. Plates are made with 0.8% phytoagar and half strength Murashige and Skoog (MS) media, adjusted to a pH of 5.7 with KOH. Sterilized seeds are suspended in 0.01% agarose to allow for them to be sown with a pipette. Plates are sealed with micropore tape, loosely covered with aluminum foil to block light, and placed in a 4C refrigerator for stratification. After 3 days, the foil is removed, and plates are moved to a growth chamber under 16/8 hr light/dark cycles. In a tissue hood, individual wells of 24-well plates are filled with 1mL of MS/2 medium with 1% sucrose. Chemical HSL, inducer, and/or bacteria are added to appropriate wells. Finally, plants are carefully lifted from the agar plates with forceps and moved to individual wells such that the roots are entirely submerged. Each plate is covered with a lid, sealed with micropore tape, and returned to the growth chamber. After 24 hours, plates are taking from the growth chamber. Plants are laid on microscope slide, the roots are covered with a cover slip, and imaged with a confocal microscope.

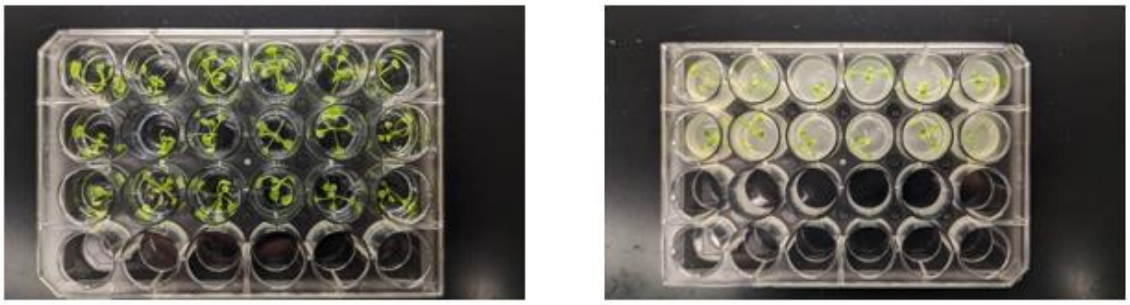

Figure 3.6 Hydroponic coculture plates.

Uncropped photographs used for Figure 3.3h which show the growth of *A. thaliana* in a hydroponic system (left) and in coculture with *P. putida* (right).

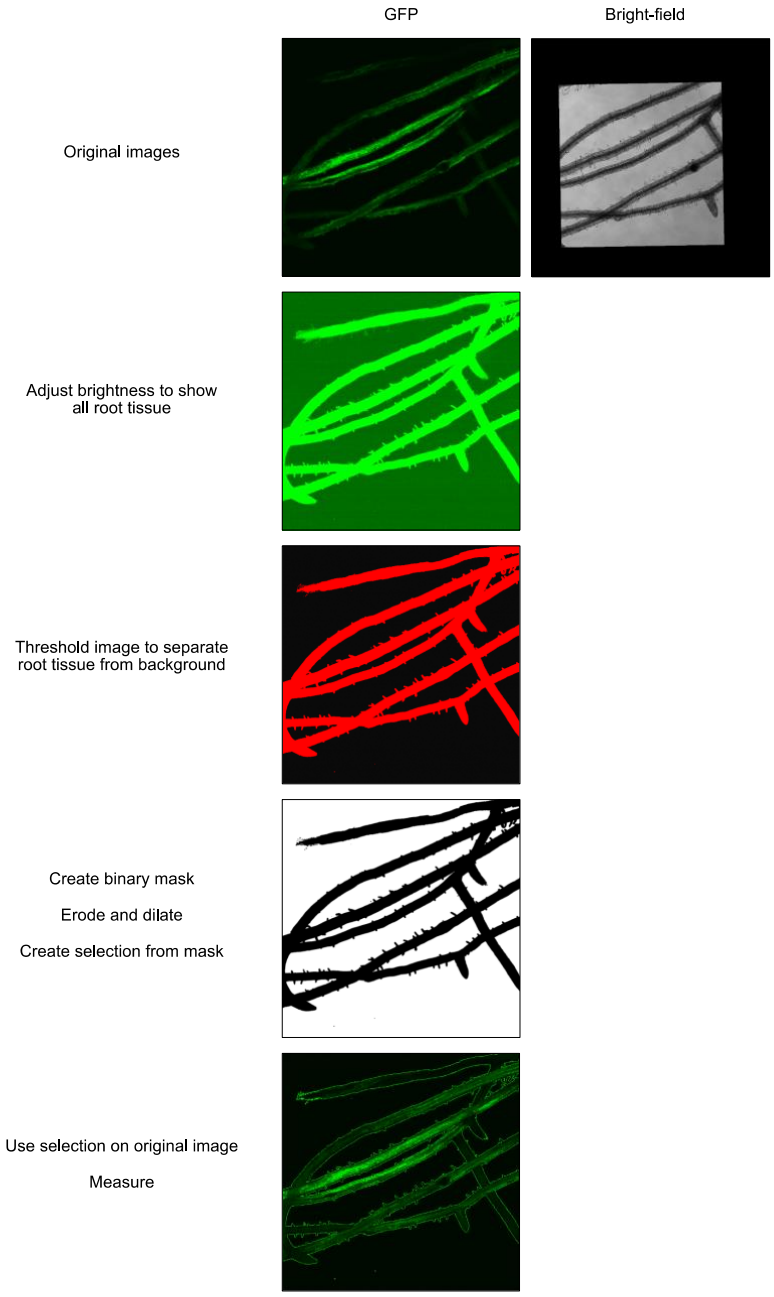

Figure 3.7 Image quantification with ImageJ.
Example images illustrating the steps of an ImageJ macro used to quantify the mean pixel intensity (MPI) of just the root tissue from a fluorescent image.

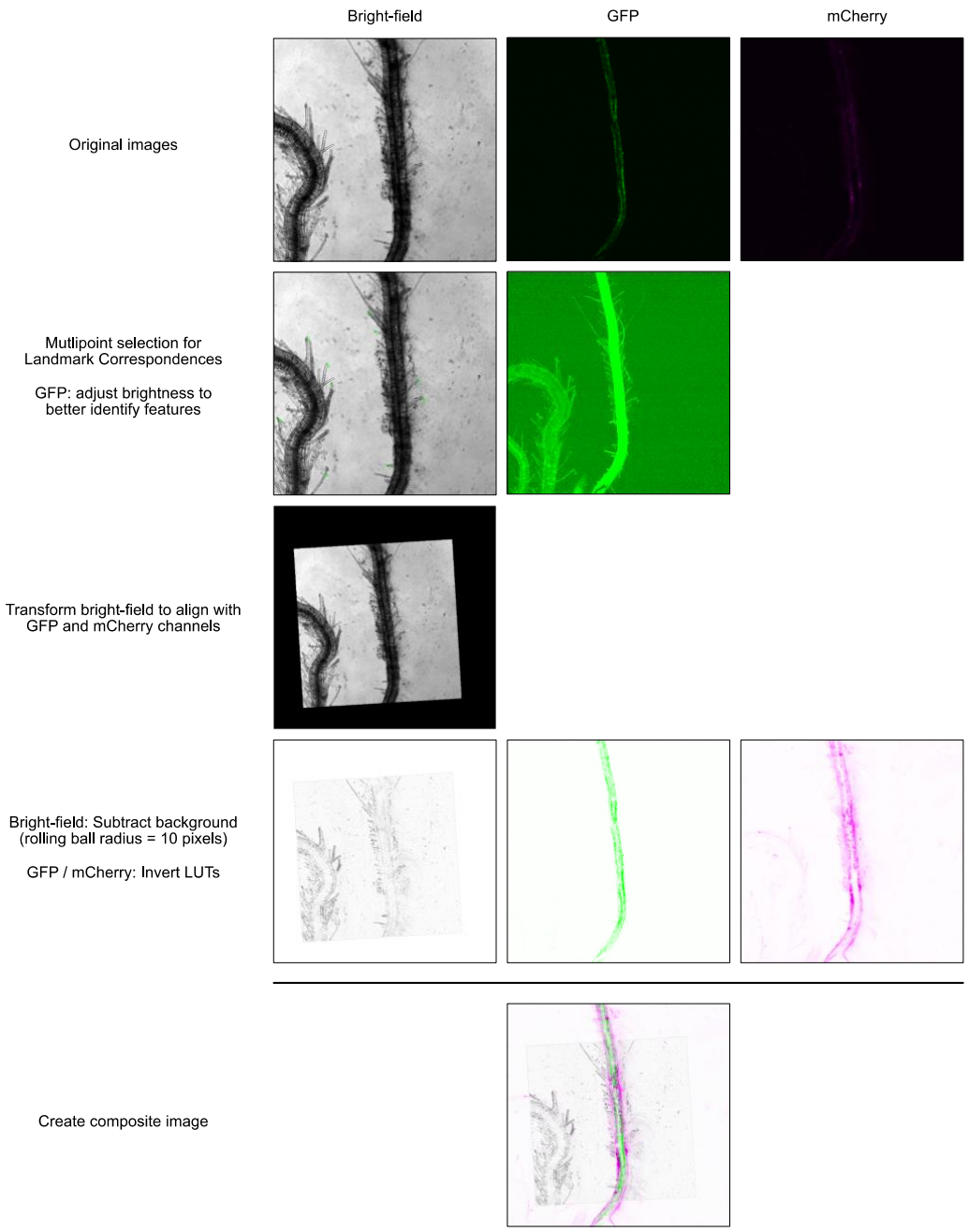

Figure 3.8 Creating composite images with ImageJ.
Example images showing the steps used to align the bright-field and fluorescent images, invert the LUTs of the fluorescent channels, and create composite images using ImageJ.

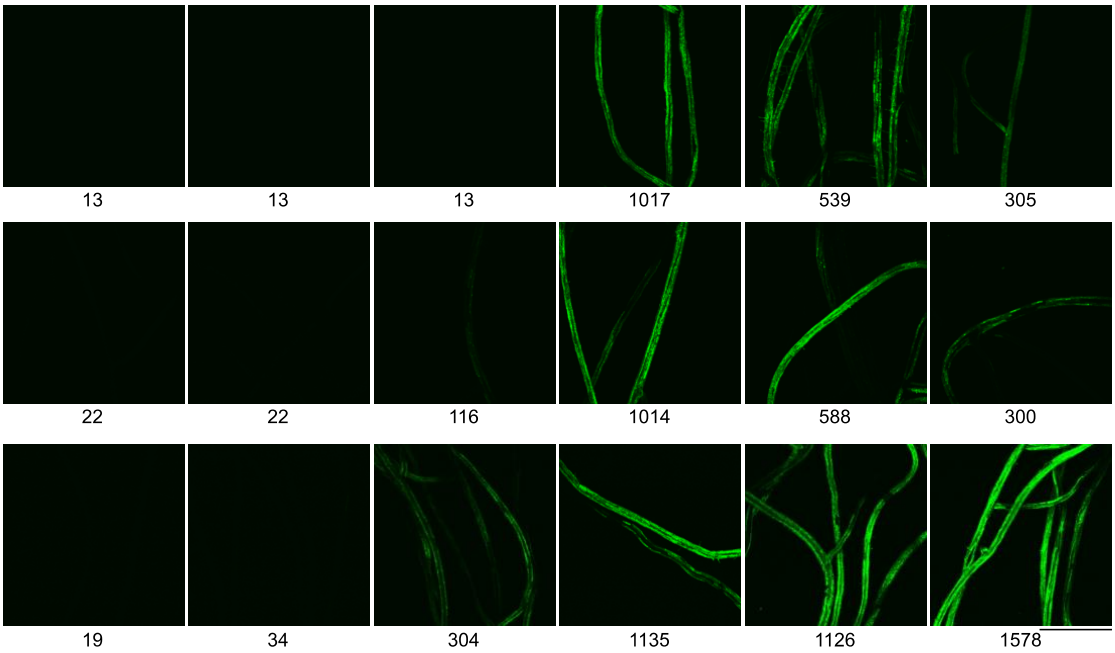

Figure 3.9 All images used for pC-HSL receiver dose response.
Raw images and measured MPI of root tissue for data plotted in Figure 3.2d and 3.2f. Scale bar represents 1mm.

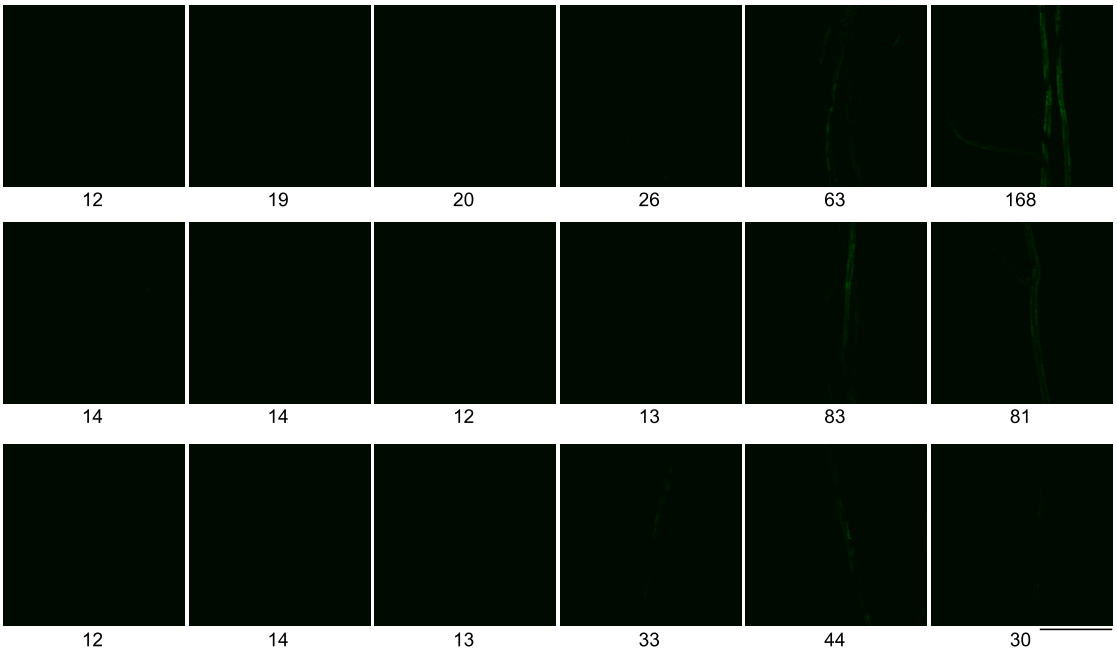

Figure 3.10 All images used for OC12 receiver dose response.
Raw images and measured MPI of root tissue for data plotted in Figure 3.2h. Scale bar represents 1mm.

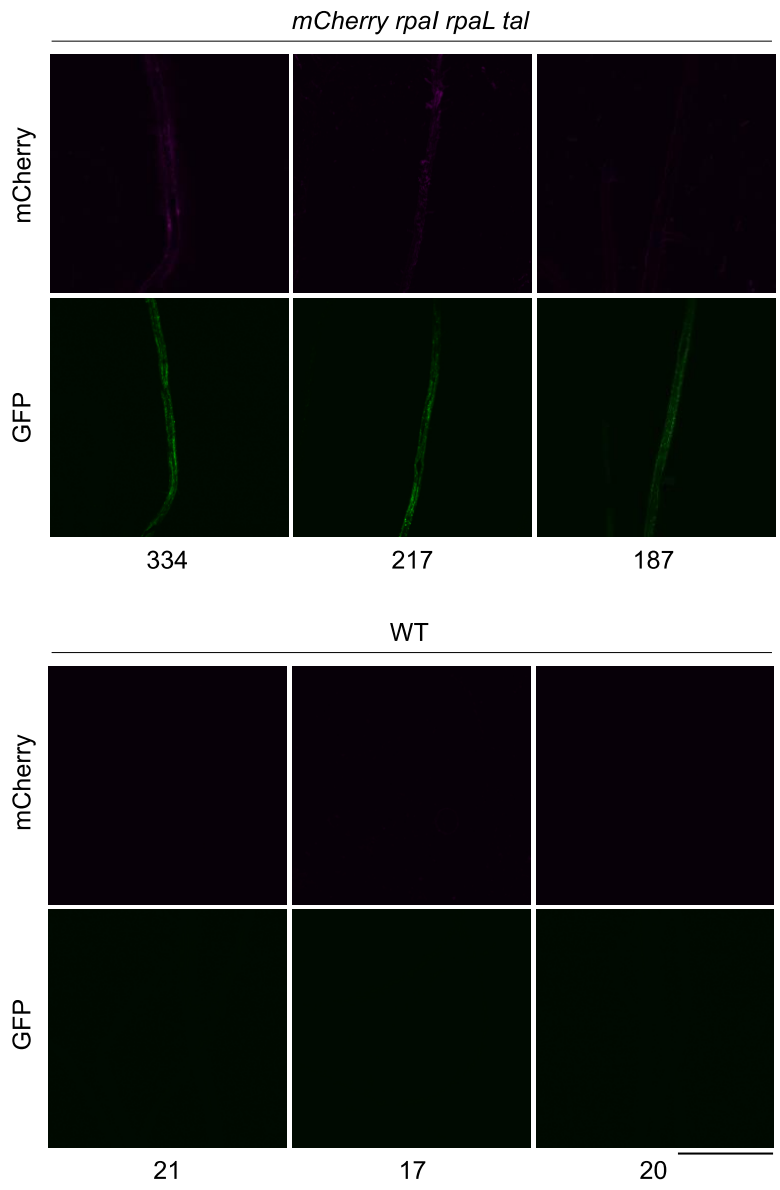

Figure 3.11 Constitutive pC-HSL sender with plants.
Raw images and measured MPI of root tissue for data plotted in Figure 3.3g. Scale bar represents 1mm.

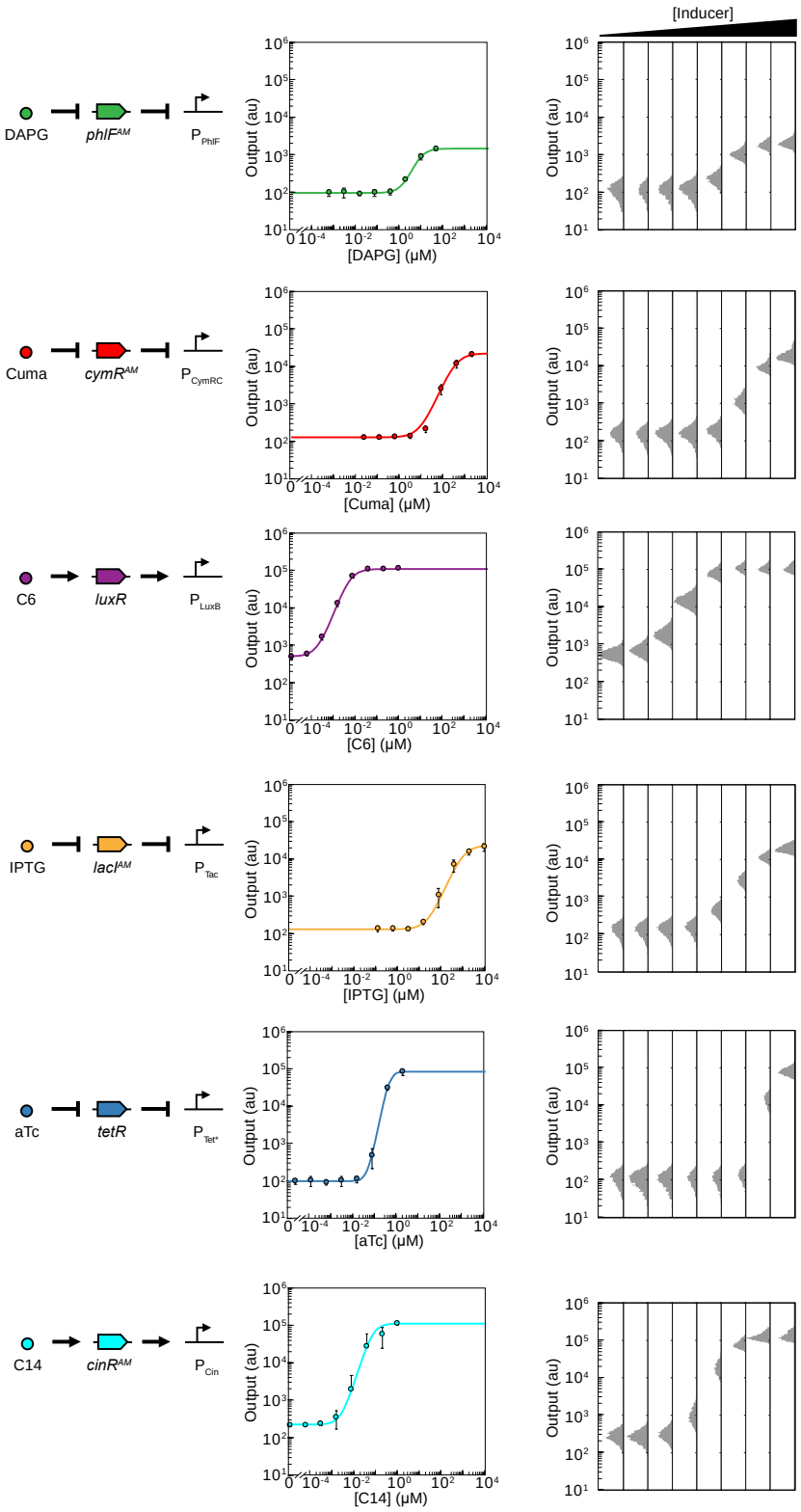

Figure 3.12 Sensor characterization in P. putida.
Response functions fit to hill equation. Table 3.1 summarizes all fit parameters.

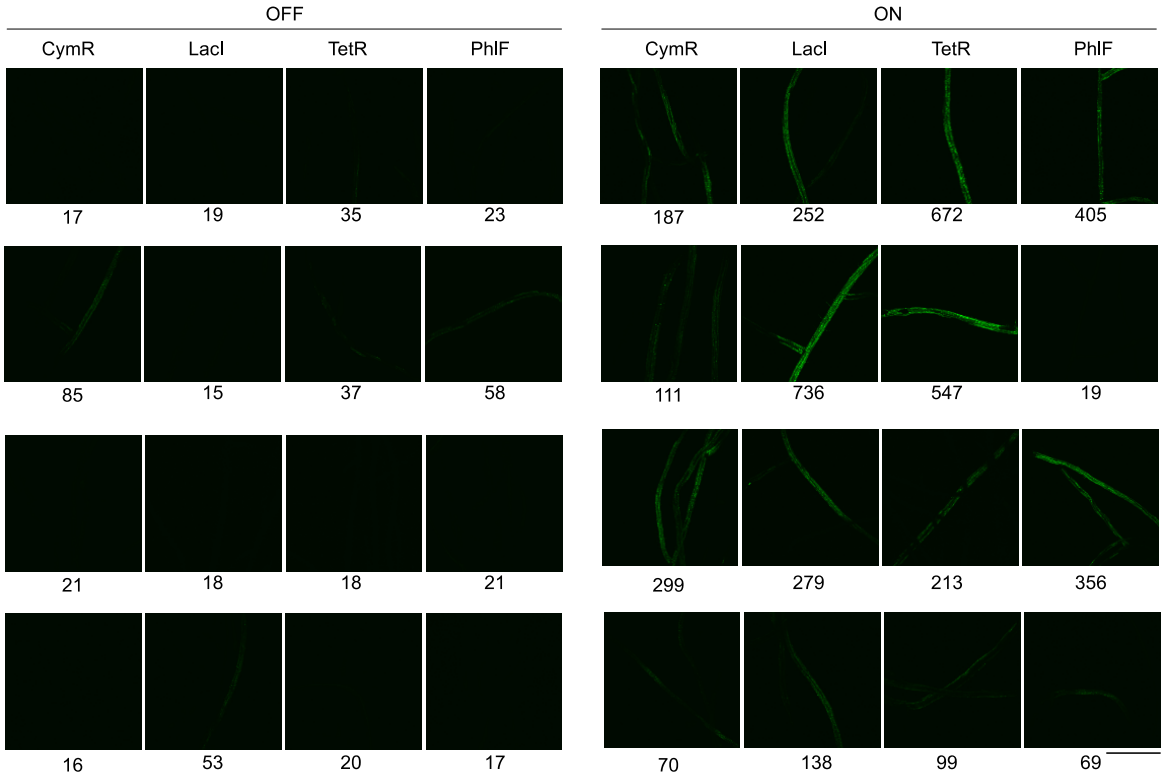

Figure 3.13 All images of sensor-sender-receiver.
Raw images and measured MPI of root tissue for data plotted in Figure 3.4d. Scale bar represents 1mm.

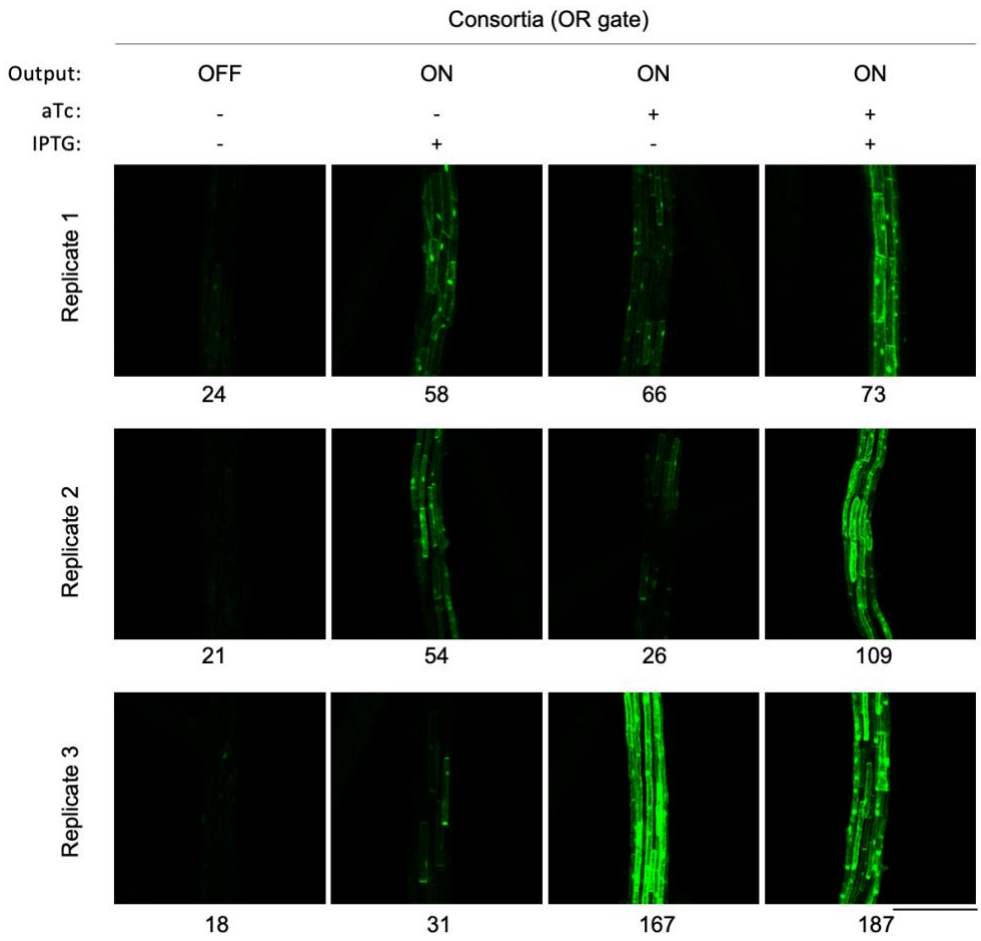

Figure 3.14 All images of consortia sensor-sender-receiver.
Cropped and brightness adjusted images and measured MPI of root tissue for data plotted in Figure 3.4g.

Scale bar represents 1mm.

Table 3.1 Response function parameters for sensors in P. putida

Sensor	Max inducer (μM)	K (μM)	*n*	y_{min} (au)	y_{max} (au)
phlF	50	7	2.0	118	1594
cymR	2000	289	1.5	135	24491
lacI	10	950	1.5	129	18715
tetR	2	0.50	3.7	117	96110

Table 3.2 Chimeric HSL sensor sequences

Sensor	DNA sequence
rpaR^{AM2}	ATGCCAAAGAAGAAGAGAAAGGTTGCCCCCCCGACCGATGTCAGCCTGGGGGACGAGCTCCACT
	TAGACGGCGAGGACGTGGCGATGGCGCATGCCGACGCGCTAGACGATTTCGATCTGGACATGTT
	GGGGGACGGGGATTCCCCGGGGCCGGGATTTACCCCCCACGACTCCGCCCCCTACGGCGCTCTG
	GATATGGCCGACTTCGAGTTTGAGCAGATGTTTACCGATGCCCTTGGAATTGACGAGTACGGTG
	GGGGAGGTGGAGGTTCTATGATTGTGGGTGAAGATCAGCTGTGGGGTCGTCGTACACTGGAATT
	TGTTGATAGCGTTGAACGTCTGGAAGCACCGGCACTGATTAGCCGTTTTGAAAGCCTGATTGCA
	AGCTGTGGTTTTACCGCCTATATCATGGCAGGTCTGCCGAGCCGTAATGCCGGTCTGCCGGAAC
	TGACCCTGGCAAATGGTTGGCCTCGTGATTGGTTTGATCTGTATGTTAGCGAAAACTTTAGCGC
	AGTTGATCCGGTTCCGCGTTATGGTGCAACCACTGTTCATCCGTTTGTTTGGAGTGATGCACCG
	TATGATCGTGACCGTGATCAGGCAGCACATCGTGTTATGACCCGTGCAGCAGAATTTGGTCTGG
	TTGAAGGTTATTGTATTCCGCTGCATTACGATGATGGTAGCGCAGCAATTAGTATGGCAGGTGA
	AGATCCTGATCTGAGTCCGGCAGCCCGTGGTGTAATGCAGCTGGTTAGCATTTATGCACATAGC
	CGTCTGCGTGTACTGAGCCGTCCGAAACCGATTCGTCGTAATCGTCTGACACCGCGTGAATGTG
	AAATTCTGCAGTGGGCAGCACAGGGTAAAACCGCATGGGAAATTAGCGTTATTCTGTGTATTAC
	CGAACGCACCGTTAAATTTCATCTGATTGAAGCAGCACGTAAACTGGATGCAGCAAATCGTACC
	GCAGCAGTTGCAAAAGCACTGACACTGGGTCTGATTCGTCTGTGA
lasR^{AM}	ATGCCAAAGAAGAAGAGAAAGGTTGCCCCCCCGACCGATGTCAGCCTGGGGGACGAGCTCCACT
	TAGACGGCGAGGACGTGGCGATGGCGCATGCCGACGCGCTAGACGATTTCGATCTGGACATGTT
	GGGGGACGGGGATTCCCCGGGGCCGGGATTTACCCCCCACGACTCCGCCCCCTACGGCGCTCTG
	GATATGGCCGACTTCGAGTTTGAGCAGATGTTTACCGATGCCCTTGGAATTGACGAGTACGGTG
	GGGGAGGTGGAGGTTCTATGGCCTTGGTTGACGGTTTTCTCGAGCTGGAACGCTCAAGTGGAAA
	ATTGGAGTTGAGCGCCATCCTGCAGAAGATGGCGAGCGACCTCGGATTCTCGAAGATCCTGTTC
	GGCCTGTTGCCTAAGGACAGCCAGGACTACGAGAACGCCTTCATCGTCAGCAACTACCCGGCCG
	CCTGGCGCGAGCATTACGACCGAACTGGCTACGCGCGGGTCGACCCGACGGTCAGTCACTGTAC
	CCAGAGCGTACTGCCGATTTTCTGGGAACCGCCCATCTACCAGACGCGAAAGCAGCACGAGTTC
	TTCGAGGAAGCCTTGGCCGCTGGCCTGGTGTATGGACTGACCATGCCGCTGCATGGTGCTCGCG
	GCGAACTCGGCGCGCTGAGCCTCAGCGTGGAAGCGGAAAACCGGGCCGAGGCCAACCGTTTCAT
	GGAGTCGGTCCTGCCGACCCTGTGGATGCTCAAGGACTACGCACTGCAGAGCGGTGTCGGAATG
	GCCTTCGAACATCCGATCAGCAAACCGGTGGTTCTGACCAGCCGAGAGAAGGAAGTGTTACAGT
	GGTGCGCCATCGGCAAGACCAGCTGGGAGATATCGGTTATCTGCAACTGCTCGGAAGCCAATGT
	GAACTTCCATATGGGAAATATTCGGCGGAAGTTCGGTGTGACCTCCCGCCGCGTAGCGGCCATT
	ATGGCCGTTAATTTGGGTCTTATTACTCTCTGA

luxR^{AM} ATGCCAAAGAAGAAGAGAAAGGTTGCCCCCCCGACCGATGTCAGCCTGGGGGACGAGCTCCACT
TAGACGGCGAGGACGTGGCGATGGCGCATGCCGACGCGCTAGACGATTTCGATCTGGACATGTT
GGGGGACGGGGATTCCCCGGGGCCGGGATTTACCCCCCACGACTCCGCCCCCTACGGCGCTCTG
GATATGGCCGACTTCGAGTTTGAGCAGATGTTTACCGATGCCCTTGGAATTGACGAGTACGGTG
GGGGAGGTGGAGGTTCTATGAAAAACATAAATGCCGACGACACATACAGAATAATTAATAAAAT
TAAAGCTTTTAGAAGCAATAATGATATTAATCAATGCTTATCTGATATGACTAAAATGGTACAT
TGTGAATATTATTTACTCGCGATCATTTATCCTCATTCTATGGTTAAATCTGATATTTCAATCC
TAGATAATTACCCTAAAAAATGGAGGCAATATTATGATGACGCTAATTTAATAAAATATGATCC
TATAGTAGATTATTCTAACTCCAATCATTCACCAATTAATTGGAATATATTTGAAAACAATGCT
GTAAATAAAAAATCTCCAAATGTAATTAAAGAAGCGAAAACATCAGGTCTTATCACTGGGTTTA
GTTTCCCTATTCATATGGCTAACAATGGCTTCGGAATGCTTAGTTTTGCATATTCAGAAAAAGA
CAACTATATAGATAGTTTATTTTTACATGCGTGTATGAACATACCATTAATTGTTCCTTCTCTA
GTTGATAATTATCGAAAAATAAATATAGCAAATAATAAATCAAACAACGATTTAACCAAAAGAG
AAAAAGAATGTTTAGCGTGGGCATGCGAAGGAAAAAGCTCTTGGGATATTTCAAAAATATTAGG
ATGCAGTGAGCGTACTGTCACTTTCCATTTAACCAATGCGCAAATGAAACTCAATACAACAAAC
CGCTGCCAAAGTATTTCTAAAGCAATTTTAACAGGAGCAATTGATTGCCCATACTTTAAAAATT
GA

cinR^{AM2} ATGCCAAAGAAGAAGAGAAAGGTTGCCCCCCCGACCGATGTCAGCCTGGGGGACGAGCTCCACT
TAGACGGCGAGGACGTGGCGATGGCGCATGCCGACGCGCTAGACGATTTCGATCTGGACATGTT
GGGGGACGGGGATTCCCCGGGGCCGGGATTTACCCCCCACGACTCCGCCCCCTACGGCGCTCTG
GATATGGCCGACTTCGAGTTTGAGCAGATGTTTACCGATGCCCTTGGAATTGACGAGTACGGTG
GGGGAGGTGGAGGTTCTATGATTGAGAATACCTATAGCGAAAAGTTCGAGTCCGCGTTCGAACA
GATCAAAGCGGCGGCCAACGTGGATGCCGCCATCCGTATTCTCCAGGCGGAATATAACCTCGAT
TTCGTCACCTACCATCTCGCCCAGACAATCGCGAGCAAGATCGATTCGCCCTTCGTGCGCACCA
CCTATCCGGATGCCTGGGTTTCCCGTTACCTCCTCAACTGCTATGTGAAGGTCGATCCGATCAT
CAAGCAGGGCTTCGAACGCCAGCTGCCCTTCGACTGGAGCGAGGTCGAACCGACGCCGGAGGCC
TATGCCATGCTGGTCGACGCCCAGAAACACGGCATCGATGACAATGGCTACTCCATCCCCGTCG
CCGACAAGGCGCAGCGCCGCGCCCTGCTGTCGCTGAATGCCCATATACCGGCCGACGAATGGAC
CGAGCTCGTGCGCCGCTACCGCAACGAGTGGATCGAGATCGCCCATCTGATCCACCGCAAGGCC
GTATATGAGCTGCATGGCGAAAACGATCCAGTGCCGGCATTGTCGCCGCGCGAGATCGAGTGTC
TGCACTGGACCGCCCTCGGCAAGGATTACAAGGATATTTCGGTCATCCTGGGCATATCAGAGCA
TACCACACGCGATTACCTGAAAACCGCCCGCTTCAGGCTCGGCTGCACCACGATCTCGGCCGCC
GCGTCGCGGGCTGTTCAATTGTGCATCATCAATCCCTATAGGATCCGCATGACGCGACGTAATT
GGTAA

Table 3.3 Synthetic HSL promoter sequences.

Regulator	Promoter	DNA sequence [a]
rpaR	P_4(rpaO):m35S	GTCTACTCCAAAAATGTCAAAGATACAGTCTCAGAAGACCAAAGGGCTAT **ACCTGTCCGATCGGACAATAACCTGTCCGATCGGACAATA**C**ACCTGTCCG ATCGGACAATA**C**ACCTGTCCGATCGGACAATA**CGCAAGACCCTTCCTCTA TATAAGGAAGTTCATTTCATTTGGAGAGGACAACCTCGAGTATTTTTACA ACAATTACCAACAACAACAAACAACAAACAACATTACAATTACTATTTAC AATTACACC
lasR	P_4(lasO):m35S	GTCTACTCCAAAAATGTCAAAGATACAGTCTCAGAAGACCAAAGGGCTAT **AACTAGCAAATGAGATAGAT**C**AACTAGCAAATGAGATAGAT**C**AACTAGCA AATGAGATAGAT**C**AACTAGCAAATGAGATAGAT**CGCAAGACCCTTCCTCT ATATAAGGAAGTTCATTTCATTTGGAGAGGACAACCTCGAGTATTTTTAC AACAATTACCAACAACAACAAACAACAAACAACATTACAATTACTATTTA CAATTACACC
luxR	P_4(luxO):m35S	GTCTACTCCAAAAATGTCAAAGATACAGTCTCAGAAGACCAAAGGGCTAT **ACCTGTAGGATCTTACAAGT**C**ACCTGTAGGATCTTACAAGT**C**ACCTGTAG GATCTTACAAGT**C**ACCTGTAGGATCTTACAAGT**CGCAAGACCCTTCCTCT ATATAAGGAAGTTCATTTCATTTGGAGAGGACAACCTCGAGTATTTTTAC AACAATTACCAACAACAACAAACAACAAACAACATTACAATTACTATTTA CAATTACACC
cinR	P_3(cinO):m35S	GTCTACTCCAAAAATGTCAAAGATACAGTCTCAGAAGACCAAAGGGCTAT **GGGGGGGCCTATCTGAGGGAAGGGGGGGCCTATCTGAGGGAAGGGGGGGC CTATCTGAGGGAA**GCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCAT TTGGAGAGGACAACCTCGAGTATTTTTACAACAATTACCAACAACAACAA ACAACAAACAACATTACAATTACTATTTACAATTACACC

a) Operator sequence – bold; m35S promoter - underlined.

Table 3.4 Plant parts sequences.

Part	Type	DNA sequence
CaMV35S	Constitutive promoter	TGAGACTTTTCAACAAAGGATAATTTCGGGAAACCTCCTCGGATTCCAT TGCCCAGCTATCTGTCACTTCATCGAAAGGACAGTAGAAAAGGAAGGTG GCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCTATCATTCAAGA TCTCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGC ATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATT GATGTGACATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCC TTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGG ACA
TMVΩ	Translational enhancer	ACAATTACCAACAACAACAAACAACAAACAACATTACAATTACTATTTA CAATTAC
SV40	Nuclear localization signal	CCAAAGAAGAAGAGAAAGGTT
VP16	Trascriptional activating domain	GCCCCCCCGACCGATGTCAGCCTGGGGGACGAGCTCCACTTAGACGGCG AGGACGTGGCGATGGCGCATGCCGACGCGCTAGACGATTTCGATCTGGA CATGTTGGGGGACGGGGATTCCCCGGGGCCGGGATTTACCCCCCACGAC TCCGCCCCCTACGGCGCTCTGGATATGGCCGACTTCGAGTTTGAGCAGA TGTTTACCGATGCCCTTGGAATTGACGAGTACGGTGGG
T_OCS	Terminator	CTGCTTTAATGAGATATGCGAGACGCCTATGATCGCATGATATTTGCTT TCAATTCTGTTGTGCACGTTGTAAAAAACCTGAGCATGTGTAGCTCAGA TCCTTACCGCCGGTTTCGGTTCATTCTAATGAATATATCACCCGTTACT ATCGTATTTTTATGAATAATATTCTCCGTTCAATTTACTGATTGTACCC TACTACTTATATGTACAATATTAAAATGAAAACAATATATTGTGCTGAA TAGGTTTATAGCGACATCTATGATAGAGCGCCACAATAACAAACAATTG CGTTTTATTATTACAAATCCAATTTTAAAAAAAGCGGCAGAACCGGTCA AACCTAAAAGACTGATTACATAAATCTTATTCAAATTTCAAAAGTGCCC CAGGGGCTAGTATCTACGACACACCGAGCGGCGAACTAATAACGCTCAC TGAAGGGAACTCCGGTTCCCCGCCGGCGCGCATGGGTGAGATTCCTTGA AGTTGAGTATTGGCCGTCCGCTCTACCGAAAGTTACGGGCACCATTCAA CCCGGTCCAGCACGGCGGCCGGGTAACCGACTTGCTGCCCCGAGAATTA TGCAGCATTTTTTTGGTGTATGTGGGCCCCAAATGAAGTGCAGGTCAAA CCTTGACAGTGACGACAAATCGTTGGGCGGGTCCAGGGCGAATTTTGCG ACAACATGTCGAGGCTCAGCAG
GGGS linker	Flexible linker	GGAGGTGGAGGTTCT

m35S	Minimal promoter	GCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGACA
sfGFP	Fluorescent reporter	ATGCGTAAAGGCGAAGAGCTGTTCACTGGTGTCGTCCCTATTCTGGTGG
		AACTGGATGGTGATGTCAACGGTCATAAGTTTTCCGTGCGTGGCGAGGG
		TGAAGGTGACGCAACTAATGGTAAACTGACGCTGAAGTTCATCTGTACT
		ACTGGTAAACTGCCGGTACCTTGGCCGACTCTGGTAACGACGCTGACTT
		ATGGTGTTCAGTGCTTTGCTCGTTATCCGGACCATATGAAGCAGCATGA
		CTTCTTCAAGTCCGCCATGCCGGAAGGCTATGTGCAGGAACGCACGATT
		TCCTTTAAGGATGACGGCACGTACAAAACGCGTGCGGAAGTGAAATTTG
		AAGGCGATACCCTGGTAAACCGCATTGAGCTGAAAGGCATTGACTTTAA
		AGAAGACGGCAATATCCTGGGCCATAAGCTGGAATACAATTTTAACAGC
		CACAATGTTTACATCACCGCCGATAAACAAAAAAATGGCATTAAAGCGA
		ATTTTAAAATTCGCCACAACGTGGAGGATGGCAGCGTGCAGCTGGCTGA
		TCACTACCAGCAAAACACTCCAATCGGTGATGGTCCTGTTCTGCTGCCA
		GACAATCACTATCTGAGCACGCAAAGCGTTCTGTCTAAAGATCCGAACG
		AGAAACGCGATCATATGGTTCTGCTGGAGTTCGTAACCGCAGCGGGCAT
		CACGCATGGTATGGATGAACTGTACAAATGA

4 Conclusions & Future Directions

In Chapter 1, I review the genetic parts and regulators available for use in plants and how they will lead to crops with improved yields and higher tolerance to a changing climate. We are beginning to have options in regards to where certain traits should be placed within a plant with emerging tools for plastid engineering. Despite these advances, genetic engineering *in planta* is much more complex than what is capable in microbes which have relatively simple and rapid lifecycles. To best take advantage of our engineering capabilities in microbes, they should be viewed as an integral part of the broader agricultural system – essentially an external plastid. Attaining this view requires forms of communication between plants and microbes in order to link the entire system. Looking to nature as a guide, there are several examples of communication molecules which may be repurposed in an engineered system by building new sensors or placing their biosynthesis under synthetic regulation. However, the search for ideal signals that are completely orthogonal remains an outstanding challenge in engineering a holistic agricultural system.

In Chapter 2, I highlight portion of a publication pertaining to the use of plant-to-microbe communication to control the expression of engineered nitrogenase clusters. Several of the sensors are for compounds that are naturally produced by plants and exudated into the soil. The use of these sensors can help ensure that the energy-intensive expression of nitrogenase only occurs when the microbes are in close proximity to a plant. Taking it one step further, we postulate that the use of engineered exudates, such as

opines, can control nitrogenase expression and act as a carbon source for the microbes, forming a synthetic symbiotic relationship. In addition, the use of an engineered plant-to-microbe signal can act as a form of biocontainment by ensuring that the engineered microbe does not function away from the desired plants.

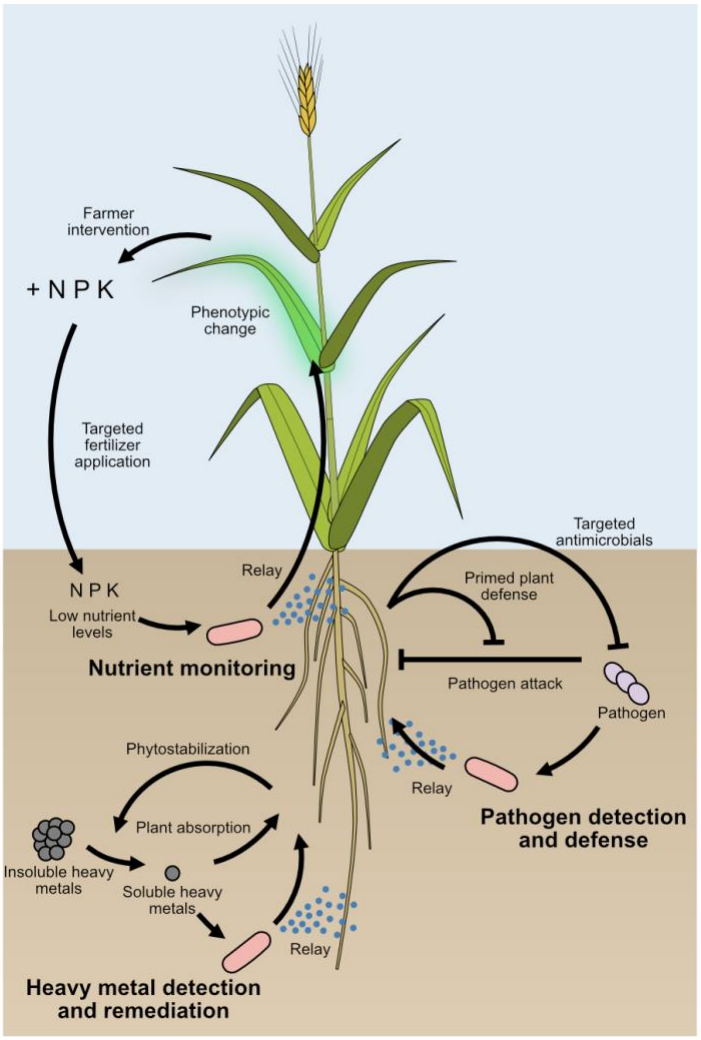

Figure 4.1 Potential applications utilizing engineered microbe-to-plant communication
Hypothetical applications of microbe-to-plant communication. *Nutrient monitoring*: Microbial sensors for nitrogen, phosphorous, and/or potassium can be used to control the production of the microbe-to-plant relay molecule in certain regimes. The relay molecule can be detected by the plant and trigger a phenotypic change, such as a change in fluorescence or pigment production, that can be detected by eye, drone, or satellite. Necessary nutrient adjustments can then be made in a targeted manner, minimizing waste and maximizing yields. *Pathogen detection and defense*: Pathogens in the soil can inhibit plant growth. Soil bacteria with engineered sensors can be used to detect specific pathogens and send a relay signal to a nearby plant. The relay signal can prime the defense system of the plant, enhancing its survival against the pathogen. In addition, the plant could produce targeted antimicrobials against the pathogen. Having dynamic control over the production of the antimicrobial minimizes the selection pressure and reduces the likelihood of the pathogen evolving resistance to the antimicrobial. *Heavy metal detection and remediation:*

Heavy metal contaminants in the soil pose a threat to human health. Soil microbes can be engineered to detect the presence of a heavy metal and relay the information to a nearby plant. Upon receiving the relay signal, the plant can produce compounds to solubilize the contaminant and begin the process of bioremediation.

In Chapter 3, I share the creation of an engineered form of microbe-to-plant communication. I engineered plants with the ability to sense and respond to bacterial quorum signals separate from native responses, thus creating a new channel of microbe-to-plant communication. In addition, we have shown that the HSL sensors are sensitive enough to respond to HSL biosynthesized by *P. putida* grown in proximity to the plants. To the best of our knowledge, this is the first demonstration of engineered bacteria to plant communication using a small molecule. We were also able to place the biosynthesis of HSLs under control of various sensors and a genetic circuit, thus creating microbial mediated phytosensors. This work lays the foundation for using bacteria grown in proximity to a plant to perform tasks such as monitoring soil nutrient content, sensing pathogens, or detecting environmental contaminants (Figure 4.1).

In summary, rather than viewing plants and microbes as separate entities, they can be viewed holistically as a single engineerable system. In this system, the desired functions can be performed in the host most suitable for the task. The distribution of work can allow for complex tasks to be performed with reduced toxicity due to resource limitations and potentially laxed regulatory concerns. To attain this goal, two-way communication between plants and microbes is essential. Ideally, synthetic forms of communication need to be as orthogonal as possible to native signaling pathways, able to diffuse through the complex soil environment, be readily produced from common cellular precursors, and easily sensed at low concentrations. While some of the molecules used in Chapters 2 and 3 might not meet all these criteria, they do provide a proof of principle for engineerable plant-microbe communication for others to build upon.

References

1. Ray, D.K., Ramankutty, N., Mueller, N.D., West, P.C. & Foley, J.A. Recent patterns of crop yield growth and stagnation. *Nat Commun* **3**, 1293 (2012).
2. Peng, S.B., Tang, Q.Y. & Zou, Y.B. Current Status and Challenges of Rice Production in China. *Plant Prod Sci* **12**, 3-8 (2009).
3. von Braun, J. Bioeconomy - The global trend and its implications for sustainability and food security. *Glob Food Secur-Agr* **19**, 81-83 (2018).
4. FAO (Rome; 2017).
5. Long, S.P., Marshall-Colon, A. & Zhu, X.-g. Meeting the Global Food Demand of the Future by Engineering Crop Photosynthesis and Yield Potential. *Cell* **161**, 56-66 (2015).
6. Francini, A. & Sebastiani, L. Abiotic Stress Effects on Performance of Horticultural Crops. *Horticulturae* **5**, 67 (2019).
7. Savary, S. et al. The global burden of pathogens and pests on major food crops. *Nature Ecology and Evolution* **3**, 430-439 (2019).
8. FAO (2019).
9. Baltes, N.J. & Voytas, D.F. Enabling plant synthetic biology through genome engineering. *Trends Biotechnol* **33**, 120-131 (2015).
10. Wurtzel, E.T. et al. Revolutionizing agriculture with synthetic biology. *Nat Plants* **5**, 1207-1210 (2019).
11. Liu, W. & Stewart, C.N. Plant synthetic biology. *Trends in Plant Science* **20**, 309-317 (2015).
12. Knott, G.J. & Doudna, J.A. CRISPR-Cas guides the future of genetic engineering. *Science* **361**, 866-869 (2018).
13. Puchta, H., Dujon, B. & Hohn, B. Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 5055-5060 (1996).
14. Lloyd, A., Plaisier, C.L., Carroll, D. & Drews, G.N. Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. *Proc Natl Acad Sci U S A* **102**, 2232-2237 (2005).
15. Mahfouz, M.M. et al. De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc Natl Acad Sci U S A* **108**, 2623-2628 (2011).
16. Shan, Q. et al. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol* **31**, 686-688 (2013).
17. Baltes, N.J. et al. Conferring resistance to geminiviruses with the CRISPR–Cas prokaryotic immune system. *Nature Plants* **1**, 1-4 (2015).
18. Reiss, B. et al. RecA stimulates sister chromatid exchange and the fidelity of double-strand break repair, but not gene targeting, in plants transformed by Agrobacterium. *Proc Natl Acad Sci U S A* **97**, 3358-3363 (2000).
19. Gao, H. et al. Heritable targeted mutagenesis in maize using a designed endonuclease. *Plant J* **61**, 176-187 (2010).
20. D'Halluin, K. et al. Targeted molecular trait stacking in cotton through targeted double-strand break induction. *Plant Biotechnology Journal* **11**, 933-941 (2013).
21. Djukanovic, V. et al. Male-sterile maize plants produced by targeted mutagenesis of the cytochrome P450-like gene (MS26) using a re-designed I-CreI homing endonuclease. *Plant J* **76**, 888-899 (2013).
22. Ainley, W.M. et al. Trait stacking via targeted genome editing. *Plant Biotechnol J* **11**, 1126-1134 (2013).
23. Christian, M. et al. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757-761 (2010).

24. Cermak, T. et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* **39**, e82 (2011).
25. Li, J.F. et al. Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nat Biotechnol* **31**, 688-691 (2013).
26. Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D. & Kamoun, S. Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**, 691-693 (2013).
27. Menz, J., Modrzejewski, D., Hartung, F., Wilhelm, R. & Sprink, T. Genome Edited Crops Touch the Market: A View on the Global Development and Regulatory Environment. *Front Plant Sci* **11**, 586027 (2020).
28. Pausch, P. et al. CRISPR-Cas Phi from huge phages is a hypercompact genome editor. *Science* **369**, 333-+ (2020).
29. Jia, H.G., Orbovic, V. & Wang, N. CRISPR-LbCas12a-mediated modification of citrus. *Plant Biotechnology Journal* **17**, 1928-1937 (2019).
30. Sheehy, R.E., Kramer, M. & Hiatt, W.R. Reduction of polygalacturonase activity in tomato fruit by antisense RNA. *Proc Natl Acad Sci U S A* **85**, 8805-8809 (1988).
31. Wang, D. et al. Characterization of CRISPR Mutants Targeting Genes Modulating Pectin Degradation in Ripening Tomato. *Plant Physiol* **179**, 544-557 (2019).
32. Waltz, E. CRISPR-edited crops free to enter market, skip regulation. *Nat Biotechnol* **34**, 582 (2016).
33. Haun, W. et al. Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. *Plant Biotechnol J* **12**, 934-940 (2014).
34. Kim, J. & Kim, J.S. Bypassing GMO regulations with CRISPR gene editing. *Nat Biotechnol* **34**, 1014-1015 (2016).
35. Waltz, E. With a free pass, CRISPR-edited plants reach market in record time. *Nat Biotechnol* **36**, 6-7 (2018).
36. Liu, L. et al. Enhancing grain-yield-related traits by CRISPR–Cas9 promoter editing of maize CLE genes. *Nature Plants* (2021).
37. Liu, X. et al. Synthesis of Seed-Specific Bidirectional Promoters for Metabolic Engineering of Anthocyanin-Rich Maize. *Plant Cell Physiol* **59**, 1942-1955 (2018).
38. Rodríguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E. & Lippman, Z.B. Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell* **171**, 470-480.e478 (2017).
39. Sauer, N.J. et al. Oligonucleotide-directed mutagenesis for precision gene editing. *Plant Biotechnol J* **14**, 496-502 (2016).
40. Anzalone, A.V., Koblan, L.W. & Liu, D.R. Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat Biotechnol* **38**, 824-844 (2020).
41. Shimatani, Z. et al. Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat Biotechnol* **35**, 441-443 (2017).
42. Zong, Y. et al. Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. *Nature Biotechnology* **36**, 950-+ (2018).
43. Tian, S. et al. Engineering herbicide-resistant watermelon variety through CRISPR/Cas9-mediated base-editing. *Plant Cell Rep* **37**, 1353-1356 (2018).
44. Svitashev, S., Schwartz, C., Lenderts, B., Young, J.K. & Mark Cigan, A. Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. *Nat Commun* **7**, 13274 (2016).
45. Li, J. et al. Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. *Nat Plants* **2**, 16139 (2016).
46. Hummel, A.W. et al. Allele exchange at the EPSPS locus confers glyphosate tolerance in cassava. *Plant Biotechnol J* **16**, 1275-1282 (2018).
47. Yu, Q.H. et al. CRISPR/Cas9-induced Targeted Mutagenesis and Gene Replacement to Generate Long-shelf Life Tomato Lines. *Sci Rep* **7**, 11874 (2017).

48. Li, X. et al. Lycopene Is Enriched in Tomato Fruit by CRISPR/Cas9-Mediated Multiplex Genome Editing. *Front Plant Sci* **9**, 559 (2018).
49. Tiwari, J.K. et al. CRISPR/Cas Genome Editing in Potato: Current Status and Future Perspectives. *Front Genet* **13**, 827808 (2022).
50. Hanania, U. et al. Establishment of a tobacco BY2 cell line devoid of plant-specific xylose and fucose as a platform for the production of biotherapeutic proteins. *Plant Biotechnol J* **15**, 1120-1129 (2017).
51. Zambryski, P. et al. Ti-Plasmid Vector for the Introduction of DNA into Plant-Cells without Alteration of Their Normal Regeneration Capacity. *Embo J* **2**, 2143-2150 (1983).
52. Vaeck, M. et al. Transgenic Plants Protected from Insect Attack. *Nature* **328**, 33-37 (1987).
53. Lemaux, P.G. Genetically Engineered Plants and Foods: A Scientist's Analysis of the Issues (Part I). *Annu Rev Plant Biol* **59**, 771-812 (2008).
54. Christou, P. Plant genetic engineering and agricultural biotechnology 1983-2013. *Trends Biotechnol* **31**, 125-127 (2013).
55. ISAAA http://www.isaaa.org/gmapprovaldatabase/event/default.asp?EventID=547. (Accessed: Sep.13, 2020).
56. Shi, J. et al. ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnol J* **15**, 207-216 (2017).
57. Butler, N.M., Baltes, N.J., Voytas, D.F. & Douches, D.S. Geminivirus-Mediated Genome Editing in Potato (Solanum tuberosum L.) Using Sequence-Specific Nucleases. *Front Plant Sci* **7**, 1045 (2016).
58. Cermak, T., Baltes, N.J., Cegan, R., Zhang, Y. & Voytas, D.F. High-frequency, precise modification of the tomato genome. *Genome Biol* **16**, 232 (2015).
59. Gil-Humanes, J. et al. High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *Plant J* **89**, 1251-1262 (2017).
60. Deng, Y. et al. Epigenetic regulation of antagonistic receptors confers rice blast resistance with yield balance. *Science* **355**, 962-965 (2017).
61. Lu, Y. et al. Targeted, efficient sequence insertion and replacement in rice. *Nat Biotechnol* **38**, 1402-1407 (2020).
62. Wang, M. et al. Gene Targeting by Homology-Directed Repair in Rice Using a Geminivirus-Based CRISPR/Cas9 System. *Mol Plant* **10**, 1007-1010 (2017).
63. Miki, D., Zhang, W., Zeng, W., Feng, Z. & Zhu, J.K. CRISPR/Cas9-mediated gene targeting in Arabidopsis using sequential transformation. *Nat Commun* **9**, 1967 (2018).
64. Strecker, J. et al. RNA-guided DNA insertion with CRISPR-associated transposases. *Science* **365**, 48-53 (2019).
65. Klompe, S.E., Vo, P.L.H., Halpin-Healy, T.S. & Sternberg, S.H. Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. *Nature* **571**, 219-225 (2019).
66. Vo, P.L.H. et al. CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering. *Nat Biotechnol* (2020).
67. Gaidukov, L. et al. A multi-landing pad DNA integration platform for mammalian cell engineering. *Nucleic Acids Res* **46**, 4072-4086 (2018).
68. Chen, Y. et al. Genetic circuit design automation for yeast. *Nat Microbiol* (2020).
69. Park, Y., Espah Borujeni, A., Gorochowski, T.E., Shin, J. & Voigt, C.A. Precision design of stable genetic circuits carried in highly-insulated E. coli genomic landing pads. *Mol Syst Biol* **16**, e9584 (2020).
70. Matzat, L.H. & Lei, E.P. Surviving an identity crisis: a revised view of chromatin insulators in the genomics era. *Biochim Biophys Acta* **1839**, 203-214 (2014).
71. Murray, A.W. & Szostak, J.W. Construction of artificial chromosomes in yeast. *Nature* **305**, 189-193 (1983).
72. Burke, D.T., Carle, G.F. & Olson, M.V. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**, 806-812 (1987).

73. Harrington, J.J., Van Bokkelen, G., Mays, R.W., Gustashaw, K. & Willard, H.F. Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat Genet* **15**, 345-355 (1997).
74. Ikeno, M. et al. Construction of YAC-based mammalian artificial chromosomes. *Nat Biotechnol* **16**, 431-439 (1998).
75. Birchler, J.A. & Han, F. Maize centromeres: structure, function, epigenetics. *Annu Rev Genet* **43**, 287-303 (2009).
76. Gaeta, R.T., Masonbrink, R.E., Krishnaswamy, L., Zhao, C. & Birchler, J.A. Synthetic chromosome platforms in plants. *Annu Rev Plant Biol* **63**, 307-330 (2012).
77. Yu, W., Lamb, J.C., Han, F. & Birchler, J.A. Telomere-mediated chromosomal truncation in maize. *Proc Natl Acad Sci U S A* **103**, 17331-17336 (2006).
78. Yu, W., Han, F., Gao, Z., Vega, J.M. & Birchler, J.A. Construction and behavior of engineered minichromosomes in maize. *Proc Natl Acad Sci U S A* **104**, 8924-8929 (2007).
79. Gaeta, R.T. et al. Recovery of a telomere-truncated chromosome via a compensating translocation in maize. *Genome* **54**, 184-195 (2011).
80. Xu, C., Cheng, Z. & Yu, W. Construction of rice mini-chromosomes by telomere-mediated chromosomal truncation. *Plant J* **70**, 1070-1079 (2012).
81. Yan, X. et al. Induction of telomere-mediated chromosomal truncation and behavior of truncated chromosomes in Brassica napus. *Plant J* **91**, 700-713 (2017).
82. Kapusi, E. et al. Telomere-mediated truncation of barley chromosomes. *Chromosoma* **121**, 181-190 (2012).
83. Yuan, J. et al. Site-specific transfer of chromosomal segments and genes in wheat engineered chromosomes. *J Genet Genomics* **44**, 531-539 (2017).
84. Nelson, A.D., Lamb, J.C., Kobrossly, P.S. & Shippen, D.E. Parameters affecting telomere-mediated chromosomal truncation in Arabidopsis. *Plant Cell* **23**, 2263-2272 (2011).
85. Teo, C.H. et al. Induction of telomere-mediated chromosomal truncation and stability of truncated chromosomes in Arabidopsis thaliana. *Plant J* **68**, 28-39 (2011).
86. Zhu, Q. et al. Development of "Purple Endosperm Rice" by Engineering Anthocyanin Biosynthesis in the Endosperm with a High-Efficiency Transgene Stacking System. *Mol Plant* **10**, 918-929 (2017).
87. Zhu, Q. et al. From Golden Rice to aSTARice: Bioengineering Astaxanthin Biosynthesis in Rice Endosperm. *Molecular Plant* **11**, 1440-1448 (2018).
88. Luo, M. et al. A five-transgene cassette confers broad-spectrum resistance to a fungal rust pathogen in wheat. *Nature Biotechnology* (2021).
89. Shih, P.M. et al. A robust gene-stacking method utilizing yeast assembly for plant synthetic biology. *Nature Communications* **7**, 1-8 (2016).
90. Hamilton, C.M., Frary, A., Lewis, C. & Tanksley, S.D. Stable transfer of intact high molecular weight DNA into plant chromosomes. *Proc Natl Acad Sci U S A* **93**, 9975-9979 (1996).
91. Abrahamian, P., Hammond, R.W. & Hammond, J. Plant Virus-Derived Vectors: Applications in Agricultural and Medical Biotechnology. *Annu Rev Virol* **7**, 513-535 (2020).
92. Born, Y. et al. The tail-associated depolymerase of Erwinia amylovora phage L1 mediates host cell adsorption and enzymatic capsule removal, which can enhance infection by other phage. *Environ Microbiol* **16**, 2168-2180 (2014).
93. McGarry, R.C., Klocko, A.L., Pang, M.X., Strauss, S.H. & Ayre, B.G. Virus-Induced Flowering: An Application of Reproductive Biology to Benefit Plant Research and Breeding. *Plant Physiol.* **173**, 47-55 (2017).
94. Bao, W., Cao, B., Zhang, Y. & Wuriyanghan, H. Silencing of Mythimna separata chitinase genes via oral delivery of in planta-expressed RNAi effectors from a recombinant plant virus. *Biotechnol Lett* **38**, 1961-1966 (2016).
95. Zhu, L. et al. Host-Induced Gene Silencing of Rice Blast Fungus Magnaporthe oryzae Pathogenicity Genes Mediated by the Brome Mosaic Virus. *Genes (Basel)* **8** (2017).

96. Hajeri, S., Killiny, N., El-Mohtar, C., Dawson, W.O. & Gowda, S. Citrus tristeza virus-based RNAi in citrus plants induces gene silencing in Diaphorina citri, a phloem-sap sucking insect vector of citrus greening disease (Huanglongbing). *J Biotechnol* **176**, 42-49 (2014).
97. Saitoh, H. et al. Production of antimicrobial defensin in Nicotiana benthamiana with a potato virus X vector. *Mol Plant Microbe Interact* **14**, 111-115 (2001).
98. Donini, M. et al. Production of an engineered killer peptide in Nicotiana benthamiana by using a potato virus X expression system. *Appl Environ Microbiol* **71**, 6360-6367 (2005).
99. Kovalskaya, N., Zhao, Y. & Hammond, R.W. Antimicrobial activity of Snakin-1 and Defensin-1 as a hybrid protein. *Phytopathology* **99**, S67-S67 (2009).
100. Shi, X.Q. et al. Efficient production of antifungal proteins in plants using a new transient expression vector derived from tobacco mosaic virus. *Plant Biotechnology Journal* **17**, 1069-1080 (2019).
101. Cody, W.B. & Scholthof, H.B. Plant Virus Vectors 3.0: Transitioning into Synthetic Genomics. *Annual Review of Phytopathology, Vol 57, 2019* **57**, 211-230 (2019).
102. Hu, J.C. et al. A barley stripe mosaic virus-based guide RNA delivery system for targeted mutagenesis in wheat and maize. *Molecular Plant Pathology* **20**, 1463-1474 (2019).
103. Tourino, A., Sanchez, E., Fereres, A. & Ponz, F. High expression of foreign proteins from a biosafe viral vector derived from Turnip mosaic virus. *Span J Agric Res* **6**, 48-58 (2008).
104. Manske, U. & Schiemann, J. Development and assessment of a potato virus X-based expression system with improved biosafety. *Environ Biosafety Res* **4**, 45-57 (2005).
105. Hernandez, C., Visser, P.B., Brown, D.J.F. & Bol, J.F. Transmission of tobacco rattle virus isolate PpK20 by its nematode vector requires one of the two non-structural genes in the viral RNA 2. *J Gen Virol* **78**, 465-467 (1997).
106. Fuentes, P., Armarego-Marriott, T. & Bock, R. Plastid transformation and its application in metabolic engineering. *Curr Opin Biotechnol* **49**, 10-15 (2018).
107. Buren, S. & Rubio, L.M. State of the art in eukaryotic nitrogenase engineering. *FEMS Microbiol Lett* **365** (2018).
108. Jarvis, P. & Lopez-Juez, E. Biogenesis and homeostasis of chloroplasts and other plastids. *Nat Rev Mol Cell Biol* **14**, 787-802 (2013).
109. Gibson, D.G. et al. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52-56 (2010).
110. Oey, M., Lohse, M., Kreikemeyer, B. & Bock, R. Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. *Plant J* **57**, 436-445 (2009).
111. Boeke, J.D. et al. GENOME ENGINEERING. The Genome Project-Write. *Science* **353**, 126-127 (2016).
112. Mellor, S.B., Behrendorff, J.B.Y.H., Nielsen, A.Z., Jensen, P.E. & Pribil, M. Non-photosynthetic plastids as hosts for metabolic engineering. *Essays Biochem* **62**, 41-50 (2018).
113. Lu, Y., Rijzaani, H., Karcher, D., Ruf, S. & Bock, R. Efficient metabolic pathway engineering in transgenic tobacco and tomato plastids with synthetic multigene operons. *Proc Natl Acad Sci U S A* **110**, E623-632 (2013).
114. Yabuta, Y. et al. Improvement of vitamin E quality and quantity in tobacco and lettuce by chloroplast genetic engineering. *Transgenic Res* **22**, 391-402 (2013).
115. Zhang, J. et al. Pest control. Full crop protection from an insect pest by expression of long double-stranded RNAs in plastids. *Science* **347**, 991-994 (2015).
116. De Marchis, F., Wang, Y., Stevanato, P., Arcioni, S. & Bellucci, M. Genetic transformation of the sugar beet plastome. *Transgenic Res* **18**, 17-30 (2009).
117. Yu, Q., Lutz, K.A. & Maliga, P. Efficient Plastid Transformation in Arabidopsis. *Plant Physiol* **175**, 186-193 (2017).
118. Ruf, S. et al. High-efficiency generation of fertile transplastomic Arabidopsis plants. *Nat Plants* **5**, 282-289 (2019).
119. Dufourmantel, N. et al. Generation of fertile transplastomic soybean. *Plant Mol Biol* **55**, 479-489 (2004).

120. Jin, S. & Daniell, H. The Engineered Chloroplast Genome Just Got Smarter. *Trends Plant Sci* **20**, 622-640 (2015).
121. Okumura, S. et al. Transformation of poplar (Populus alba) plastids and expression of foreign proteins in tree chloroplasts. *Transgenic Res* **15**, 637-646 (2006).
122. Maliga, P. Plastid transformation in higher plants. *Annu Rev Plant Biol* **55**, 289-313 (2004).
123. Khan, M.S. & Maliga, P. Fluorescent antibiotic resistance marker for tracking plastid transformation in higher plants. *Nat Biotechnol* **17**, 910-915 (1999).
124. Lee, S.M. et al. Plastid transformation in the monocotyledonous cereal crop, rice (Oryza sativa) and transmission of transgenes to their progeny. *Mol Cells* **21**, 401-410 (2006).
125. Yang, X., Jiang, J., Li, Z., Liang, J. & Xiang, Y. Strategies for mitochondrial gene editing. *Comput Struct Biotechnol J* **19**, 3319-3329 (2021).
126. Nakazato, I. et al. Targeted base editing in the plastid genome of Arabidopsis thaliana. *Nat Plants* **7**, 906-913 (2021).
127. Kang, B.C. et al. Chloroplast and mitochondrial DNA editing in plants. *Nat Plants* **7**, 899-905 (2021).
128. Min, S.R. et al. An episomal vector system for plastid transformation in higher plants. *Plant Biotechnol Rep* **9**, 443-449 (2015).
129. Staub, J.M. & Maliga, P. Extrachromosomal elements in tobacco plastids. *Proc Natl Acad Sci U S A* **91**, 7468-7472 (1994).
130. Occhialini, A. et al. Mini-synplastomes for plastid genetic engineering. *Plant Biotechnol J* **20**, 360-373 (2022).
131. Fox, T.D., Sanford, J.C. & Mcmullin, T.W. Plasmids Can Stably Transform Yeast Mitochondria Lacking Endogenous Mtdna. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 7288-7292 (1988).
132. Johnston, S.A., Anziano, P.Q., Shark, K., Sanford, J.C. & Butow, R.A. Mitochondrial Transformation in Yeast by Bombardment with Microprojectiles. *Science* **240**, 1538-1541 (1988).
133. Larosa, V., Coosemans, N., Motte, P., Bonnefoy, N. & Remacle, C. Reconstruction of a human mitochondrial complex I mutation in the unicellular green alga Chlamydomonas. *Plant J* **70**, 759-768 (2012).
134. Remacle, C., Cardol, P., Coosemans, N., Gaisne, M. & Bonnefoy, N. High-efficiency biolistic transformation of Chlamydomonas mitochondria can be used to insert mutations in complex I genes. *Proc Natl Acad Sci U S A* **103**, 4771-4776 (2006).
135. Gammage, P.A., Moraes, C.T. & Minczuk, M. Mitochondrial Genome Engineering: The Revolution May Not Be CRISPR-Ized. *Trends Genet* **34**, 101-110 (2018).
136. Mok, B.Y. et al. A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. *Nature* **583**, 631-637 (2020).
137. Farre, J.C. & Araya, A. Gene expression in isolated plant mitochondria: high fidelity of transcription, splicing and editing of a transgene product in electroporated organelles. *Nucleic Acids Res* **29**, 2484-2491 (2001).
138. Koulintchenko, M., Konstantinov, Y. & Dietrich, A. Plant mitochondria actively import DNA via the permeability transition pore complex. *Embo J* **22**, 1245-1254 (2003).
139. Buren, S., Jiang, X., Lopez-Torrejon, G., Echavarri-Erasun, C. & Rubio, L.M. Purification and In Vitro Activity of Mitochondria Targeted Nitrogenase Cofactor Maturase NifB. *Front Plant Sci* **8**, 1567 (2017).
140. Kazama, T. et al. Curing cytoplasmic male sterility via TALEN-mediated mitochondrial genome editing. *Nat Plants* **5**, 722-730 (2019).
141. Arimura, S.I. et al. Targeted gene disruption of ATP synthases 6-1 and 6-2 in the mitochondrial genome of Arabidopsis thaliana by mitoTALENs. *Plant J* **104**, 1459-1471 (2020).
142. Lee, H. et al. Mitochondrial DNA editing in mice with DddA-TALE fusion deaminases. *Nat Commun* **12**, 1190 (2021).

143. Mok, Y.G. et al. Base editing in human cells with monomeric DddA-TALE fusion deaminases. *Nature Communications* **13**, 1-10 (2022).
144. Endy, D. Foundations for engineering biology. *Nature* **438**, 449-453 (2005).
145. (
146. Kohli, A. et al. Transgene integration, organization and interaction in plants. *Plant Molecular Biology* **52**, 247-258 (2003).
147. Collier, R. et al. Accurate measurement of transgene copy number in crop plants using droplet digital PCR. *Plant Journal* **90**, 1014-1025 (2017).
148. Kosuri, S. et al. Composability of regulatory sequences controlling transcription and translation in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 14024-14029 (2013).
149. Nielsen, A.A., Segall-Shapiro, T.H. & Voigt, C.A. Advances in genetic circuit design: novel biochemistries, deep part mining, and precision gene expression. *Curr Opin Chem Biol* **17**, 878-892 (2013).
150. Hossain, A. et al. Automated design of thousands of nonrepetitive parts for engineering stable genetic systems. *Nat Biotechnol* **38**, 1466-1475 (2020).
151. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* **27**, 946-950 (2009).
152. Zhou, X., Carranco, R., Vitha, S. & Hall, T.C. The dark side of green fluorescent protein. *New Phytol* **168**, 313-322 (2005).
153. Hraška, M., Rakouský, S. & Čurn, V. Green fluorescent protein as a vital marker for non-destructive detection of transformation events in transgenic plants. *Plant cell, tissue and organ culture* **86**, 303-318 (2006).
154. Kinney, J.B., Murugan, A., Callan, C.G. & Cox, E.C. Using deep sequencing to characterize the biophysical mechanism of a transcriptional regulatory sequence. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 9158-9163 (2010).
155. Schaumberg, K.A. et al. Quantitative characterization of genetic parts and circuits for plant synthetic biology. *Nat Methods* **13**, 94-100 (2016).
156. Lin, Q.P. et al. Prime genome editing in rice and wheat. *Nature Biotechnology* **38**, 582-+ (2020).
157. Jin, S. et al. Genome-wide specificity of prime editors in plants. *Nature Biotechnology* (2021).
158. Lai, X. et al. Building Transcription Factor Binding Site Models to Understand Gene Regulation in Plants. *Molecular Plant* **12**, 743-763 (2019).
159. Smale, S.T. & Kadonaga, J.T. The RNA polymerase II core promoter. *Annu Rev Biochem* **72**, 449-479 (2003).
160. Dutt, M., Dhekney, S.A., Soriano, L., Kandel, R. & Grosser, J.W. Temporal and spatial control of gene expression in horticultural crops. *Horticulture Research* **1**, 1-17 (2014).
161. Bush, M.S. et al. Selective recruitment of proteins to 5′ cap complexes during the growth cycle in Arabidopsis. *Plant Journal* **59**, 400-412 (2009).
162. Amack, S.C. & Antunes, M.S. CaMV35S promoter – A plant biology and biotechnology workhorse in the era of synthetic biology. *Current Plant Biology* **24**, 100179-100179 (2020).
163. Kummari, D. et al. An update and perspectives on the use of promoters in plant genetic engineering. *Journal of Biosciences* **45** (2020).
164. Meyer, P. & Saedler, H. Homology-Dependent Gene Silencing in Plants. *Annu Rev Plant Physiol Plant Mol Biol* **47**, 23-48 (1996).
165. Belcher, M.S. et al. Design of orthogonal regulatory systems for modulating gene expression in plants. *Nature Chemical Biology* (2020).
166. Cai, Y.-M. et al. Rational design of minimal synthetic promoters for plants. *Nucleic Acids Research* **48**, 11845-11856 (2020).
167. Feike, D. et al. Characterizing standard genetic parts and establishing common principles for engineering legume and cereal roots. *Plant Biotechnology Journal* **17**, 2234-2245 (2019).

168. Butaye, K.M. et al. Stable high-level transgene expression in Arabidopsis thaliana using gene silencing mutants and matrix attachment regions. *Plant J* **39**, 440-449 (2004).
169. Mlynarova, L. et al. Reduced Position Effect in Mature Transgenic Plants Conferred by the Chicken Lysozyme Matrix-Associated Region. *Plant Cell* **6**, 417-426 (1994).
170. Shahmuradov, I.A., Gammerman, A.J., Hancock, J.M., Bramley, P.M. & Solovyev, V.V. PlantProm: A database of plant promoter sequences. *Nucleic Acids Research* **31**, 114-117 (2003).
171. Sun, Q. et al. PPDB, the Plant Proteomics Database at Cornell. *Nucleic Acids Research* **37**, 969-974 (2009).
172. Shahmuradov, I.A., Umarov, R.K. & Solovyev, V.V. TSSPlant: A new tool for prediction of plant Pol II promoters. *Nucleic Acids Research* **45** (2017).
173. Chen, Y.-J. et al. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. **10**, 659 (2013).
174. Cambray, G. et al. Measurement and modeling of intrinsic transcription terminators. *Nucleic Acids Res* **41**, 5139-5148 (2013).
175. Peters, J.M., Vangeloff, A.D. & Landick, R. Bacterial transcription terminators: the RNA 3'-end chronicles. *J Mol Biol* **412**, 793-813 (2011).
176. O'Sullivan, J.M. et al. Gene loops juxtapose promoters and terminators in yeast. *Nat Genet* **36**, 1014-1018 (2004).
177. Curran, K.A., Karim, A.S., Gupta, A. & Alper, H.S. Use of expression-enhancing terminators in Saccharomyces cerevisiae to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metab Eng* **19**, 88-97 (2013).
178. Andreou, A.I., Nirkko, J., Ochoa-Villarreal, M. & Nakayama, N. Mobius Assembly for Plant Systems highlights promoter-terminator interaction in gene regulation. *bioRxiv* (2021).
179. Randise-Hinchliff, C.E. & Brickner, J.H. A new direction for gene looping. *Dev Cell* **23**, 919-921 (2012).
180. Moabbi, A.M., Agarwal, N., El Kaderi, B. & Ansari, A. Role for gene looping in intron-mediated enhancement of transcription. *Proc Natl Acad Sci U S A* **109**, 8505-8510 (2012).
181. Tan-Wong, S.M. et al. Gene loops enhance transcriptional directionality. *Science* **338**, 671-675 (2012).
182. Mukundan, B. & Ansari, A. Srb5/Med18-mediated termination of transcription is dependent on gene looping. *J Biol Chem* **288**, 11384-11394 (2013).
183. Medler, S. & Ansari, A. Gene looping facilitates TFIIH kinase-mediated termination of transcription. *Sci Rep* **5**, 12586 (2015).
184. F. de Felippes, F. et al. The key role of terminators on the expression and post-transcriptional gene silencing of transgenes. *Plant Journal* **104**, 96-112 (2020).
185. Nagaya, S., Kawamura, K., Shinmyo, A. & Kato, K. The HSP terminator of arabidopsis thaliana increases gene expression in plant cells. *Plant and Cell Physiology* **51**, 328-332 (2010).
186. Merchante, C., Stepanova, A.N. & Alonso, J.M. Translation regulation in plants: an interesting past, an exciting present and a promising future. *Plant Journal* **90**, 628-653 (2017).
187. Miras, M., Allen Miller, W., Truniger, V. & Aranda, M.A. Non-canonical translation in Plant RNA viruses. *Front Plant Sci* **8** (2017).
188. Xu, G. et al. uORF-mediated translation allows engineered plant disease resistance without fitness costs. *Nature* **545**, 491-494 (2017).
189. Baird, S.D., Turcotte, M., Korneluk, R.G. & Holcik, M. Searching for IRES. *Rna* **12**, 1755-1785 (2006).
190. Dorokhov, Y.L. et al. Polypurine (A)-rich sequences promote cross-kingdom conservation of internal ribosome entry. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 5301-5306 (2002).
191. Zhu, B., Zhang, W., Zhang, T., Liu, B. & Jiang, J. Genome-wide prediction and validation of intergenic enhancers in arabidopsis using open chromatin signatures. *Plant Cell* **27**, 2415-2426 (2015).

192. Bally, J. et al. Improved insect-proofing: expressing double-stranded RNA in chloroplasts. *Pest Manag Sci* **74**, 1751-1758 (2018).
193. Kanoria, S. & Burma, P.K. A 28 nt long synthetic 5 ′ UTR (synJ) as an enhancer of transgene expression in dicotyledonous plants. *BMC Biotechnology* **12**, 1-1 (2012).
194. Le Hir, H., Nott, A. & Moore, M.J. How introns influence and enhance eukaryotic gene expression. *Trends in Biochemical Sciences* **28**, 215-220 (2003).
195. Bartlett, J.G., Snape, J.W. & Harwood, W.A. Intron-mediated enhancement as a method for increasing transgene expression levels in barley. *Plant Biotechnology Journal* **7**, 856-866 (2009).
196. Laxa, M. Intron-mediated enhancement: A tool for heterologous gene expression in plants? *Front Plant Sci* **7**, 1-13 (2017).
197. Liao, L., Ning, G., Liu, C., Zhang, W. & Bao, M. The intron from the 5'-UTR of the FBP11 gene in petunia displays promoter- and enhancer-like functions. *Scientia Horticulturae* **154**, 96-101 (2013).
198. Faden, F. et al. Phenotypes on demand via switchable target protein degradation in multicellular organisms. *Nature Communications* **7** (2016).
199. Faden, F., Mielke, S. & Dissmeyer, N. Modulating protein stability to switch toxic protein function on and off in living cells. *Plant Physiol.* **179**, 929-942 (2019).
200. Banaszynski, L.A., Chen, L.C., Maynard-Smith, L.A., Ooi, A.G. & Wandless, T.J. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell* **126**, 995-1004 (2006).
201. Bonger, K.M., Chen, L.C., Liu, C.W. & Wandless, T.J. Small-molecule displacement of a cryptic degron causes conditional protein degradation. *Nat Chem Biol* **7**, 531-537 (2011).
202. Kosugi, S. et al. Six classes of nuclear localization signals specific to different binding grooves of importinα. *J. Biol. Chem.* **284**, 478-485 (2009).
203. Patron, N.J. & Waller, R.F. Transit peptide diversity and divergence: A global analysis of plastid targeting signals. *Bioessays* **29**, 1048-1058 (2007).
204. Allen, R.S. et al. Expression of 16 Nitrogenase Proteins within the Plant Mitochondrial Matrix. *Front Plant Sci* **8** (2017).
205. Lee, S. et al. Mitochondrial targeting of the Arabidopsis F1-ATPase gamma-subunit via multiple compensatory and synergistic presequence motifs. *Plant Cell* **24**, 5037-5057 (2012).
206. Halpin, C., Cooke, S.E., Barakate, A., El Amrani, A. & Ryan, M.D. Self-processing 2A-polyproteins - A system for co-ordinate expression of multiple proteins in transgenic plants. *Plant Journal* **17**, 453-459 (1999).
207. Ha, S.H. et al. Application of two bicistronic systems involving 2A and IRES sequences to the biosynthesis of carotenoids in rice endosperm. *Plant Biotechnology Journal* **8**, 928-938 (2010).
208. Yang, D. et al. Repurposing type III polyketide synthase as a malonyl-CoA biosensor for metabolic engineering in bacteria. *PNAS* **115**, 9835-9844 (2018).
209. Gong, Y.N. et al. The NIa-Protease Protein Encoded by the Pepper Mottle Virus Is a Pathogenicity Determinant and Releases DNA Methylation of Nicotiana benthamiana. *Front Microbiol* **11**, 102 (2020).
210. Goh, C.J. & Hahn, Y. Analysis of proteolytic processing sites in potyvirus polyproteins revealed differential amino acid preferences of NIa-Pro protease in each of seven cleavage sites. *PLoS One* **16**, e0245853 (2021).
211. Gallego-Bartolomé, J. et al. Targeted DNA demethylation of the arabidopsis genome using the human TET1 catalytic domain. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E2125-E2134 (2018).
212. Veley, K.M. et al. Improving disease resistance in plants by editing the epigenome. *bioRxiv* (2022).
213. Papikian, A., Liu, W., Gallego-Bartolomé, J. & Jacobsen, S.E. Site-specific manipulation of Arabidopsis loci using CRISPR-Cas9 SunTag systems. *Nature Communications* **10**, 1-11 (2019).
214. Ratcliff, F., Harrison, B.D. & Baulcombe, D.C. A similarity between viral defense and gene silencing in plants. *Science* **276**, 1558-1560 (1997).

215. Lakatos, L., Szittya, G., Silhavy, D. & Burgyán, J. Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. *Embo J* **23**, 876-884 (2004).
216. Brophy, J.A.N. & Voigt, C.A. Principles of Genetic Circuit Design. *Nature methods* **11**, 508-520 (2014).
217. McAdams, H.H. & Arkin, A. It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet* **15**, 65-69 (1999).
218. Lou, C., Stanton, B., Chen, Y.J., Munsky, B. & Voigt, C.A. Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nat Biotechnol* **30**, 1137-1142 (2012).
219. Mutalik, V.K. et al. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat Methods* **10**, 354-360 (2013).
220. Segall-Shapiro, T.H., Sontag, E.D. & Voigt, C.A. Engineered promoters enable constant gene expression at any copy number in bacteria. *Nat Biotechnol* **36**, 352-358 (2018).
221. Kushwaha, M.a.S.H.M. A portable expression resource for engineering cross-species genetic circuits and pathways. *Nature Communications* **6** (2015).
222. Zong, Y. et al. Insulated transcriptional elements enable precise design of genetic circuits. *Nat Commun* **8**, 52 (2017).
223. Legen, J. et al. Stabilization and translation of synthetic operon-derived mRNAs in chloroplasts by sequences representing PPR protein-binding sites. *Plant J* **94**, 8-21 (2018).
224. Rojas, M., Yu, Q., Williams-Carrier, R., Maliga, P. & Barkan, A. Engineered PPR proteins as inducible switches to activate the expression of chloroplast transgenes. *Nature Plants* **5**, 505-511 (2019).
225. Yu, Q., Barkan, A. & Maliga, P. Engineered RNA-binding protein for transgene activation in non-green plastids. *Nat Plants* **5**, 486-490 (2019).
226. Yu, Q., Tungsuchat-Huang, T., Verma, K., Radler, M.R. & Maliga, P. Independent translation of ORFs in dicistronic operons, synthetic building blocks for polycistronic chloroplast gene expression. *Plant J* (2020).
227. Perez-Gonzalez, A. & Caro, E. Benefits of using genomic insulators flanking transgenes to increase expression and avoid positional effects. *Sci Rep-Uk* **9** (2019).
228. Dong, P. et al. 3D Chromatin Architecture of Large Plant Genomes Determined by Local A/B Compartments. *Molecular Plant* **10**, 1497-1509 (2017).
229. Fujita, M. et al. Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Current Opinion in Plant Biology* **9**, 436-442 (2006).
230. Zhu, X.-G., Ort, D.R., Parry, M.A.J. & von Caemmerer, S. A wish list for synthetic biology in photosynthesis research. *J Exp Bot* **71**, 2219-2225 (2020).
231. Farmer, E.E. Surface-to-air signals. *Nature* **411**, 854-856 (2001).
232. Steinbrenner, A.D. et al. A receptor-like protein mediates plant immune responses to herbivore-associated molecular patterns. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 31510-31518 (2020).
233. Tena, G., Boudsocq, M. & Sheen, J. Protein kinase signaling networks in plant innate immunity. *Current Opinion in Plant Biology* **14**, 519-529 (2011).
234. Ueda, Y. & Yanagisawa, S. Delineation of Nitrogen Signaling Networks: Computational Approaches in the Big Data Era. *Molecular Plant* **12**, 150-152 (2019).
235. Sheen, J. Master regulators in plant glucose signaling networks. *Journal of Plant Biology* **57**, 67-79 (2014).
236. Scotti, N. & Cardi, T. Transgene-induced pleiotropic effects in transplastomic plants. *Biotechnol Lett* **36**, 229-239 (2014).
237. Eom, J.-S. et al. Diagnostic kit for rice blight resistance. *Nature Biotechnology* **37**, 1372-1379 (2019).

238. Khodakovskaya, M., McAvoy, R., Peters, J., Wu, H. & Li, Y. Enhanced cold tolerance in transgenic tobacco expressing a chloroplast ω-3 fatty acid desaturase gene under the control of a cold-inducible promoter. *Planta* **223**, 1090-1100 (2006).
239. Hoshika, S. et al. Hachimoji DNA and RNA: A genetic system with eight building blocks. *Science* **363**, 884-887 (2019).
240. Basu, S., Mehreja, R., Thiberge, S., Chen, M.-T. & Weiss, R. Spatiotemporal control of gene expression with pulse-generating networks. *PNAS* **101**, 6355-6360 (2004).
241. Sohka, T. et al. An externally tunable bacterial band-pass filter. *PNAS* **106**, 10135-10140 (2009).
242. Brophy, J.A. & Voigt, C.A. Principles of genetic circuit design. *Nat Methods* **11**, 508-520 (2014).
243. Blount, B.A., Weenink, T. & Ellis, T. Construction of synthetic regulatory networks in yeast. *FEBS Lett.* **586**, 2112-2121 (2012).
244. Boni, R. et al. Pathogen-inducible Ta-Lr34res expression in heterologous barley confers disease resistance without negative pleiotropic effects. *Plant Biotechnology Journal* **16**, 245-253 (2018).
245. Yang, X. et al. Biological Parts for Plant Biodesign to Enhance Land-Based Carbon Dioxide Removal. *BioDesign Research* **2021** (2021).
246. Endo, S., Betsuyaku, S. & Fukuda, H. Endogenous peptide ligand-receptor systems for diverse signaling networks in plants. *Current Opinion in Plant Biology* **21**, 140-146 (2014).
247. Liu, S.-M. et al. Transgenic plants expressing the AaIT/GNA fusion protein show increased resistance and toxicity to both chewing and sucking pests. *Insect Science* **23**, 265-276 (2015).
248. Dóczi, R. et al. The Arabidopsis mitogen-activated protein kinase kinase MKK3 is upstream of group C mitogen-activated protein kinases and participates in pathogen signaling. *Plant Cell* **19**, 3266-3279 (2007).
249. Wang, J., Mei, J. & Ren, G. Plant microRNAs: Biogenesis, homeostasis, and degradation. *Front Plant Sci* **10**, 1-12 (2019).
250. Yilmaz, A. et al. AGRIS: The arabidopsis gene regulatory information server, an update. *Nucleic Acids Research* **39**, 1118-1122 (2011).
251. Jones, D.M. & Vandepoele, K. Identification and evolution of gene regulatory networks: insights from comparative studies in plants. *Current Opinion in Plant Biology* **54**, 42-48 (2020).
252. Haque, S., Ahmad, J.S., Clark, N.M., Williams, C.M. & Sozzani, R. Computational prediction of gene regulatory networks in plant growth and development. *Current Opinion in Plant Biology* **47**, 96-105 (2019).
253. Genoud, T., Trevino Santa Cruz, M.B. & Métraux, J.P. Numeric simulation of plant signaling networks. *Plant Physiol.* **126**, 1430-1437 (2001).
254. Tu, X. et al. Reconstructing the maize leaf regulatory network using ChIP-seq data of 104 transcription factors. *Nat Commun* **11**, 5089 (2020).
255. O'Malley, R.C. et al. Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**, 1280-1292 (2016).
256. Taylor-Teeples, M. et al. An Arabidopsis gene regulatory network for secondary cell wall synthesis. *Nature* **517**, 571-575 (2015).
257. Li, G.-W., Burkhardt, D., Gross, C. & Weissman, J.S. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157**, 624-635 (2014).
258. Ikeuchi, M. et al. A Gene Regulatory Network for Cellular Reprogramming in Plant Regeneration. *Plant Cell Physiol* **59**, 765-777 (2018).
259. Hussey, S.G. et al. A Standardized Synthetic Eucalyptus Transcription Factor and Promoter Panel for Re-engineering Secondary Cell Wall Regulation in Biomass and Bioenergy Crops. *ACS Synth Biol* **8**, 463-465 (2019).
260. Timmermann, T., Gonzalez, B. & Ruz, G.A. Reconstruction of a gene regulatory network of the induced systemic resistance defense response in Arabidopsis using boolean networks. *BMC Bioinformatics* **21**, 142 (2020).

261. Koschmann, J. et al. Integration of Bioinformatics and Synthetic Promoters Leads to the Discovery of Novel Elicitor-Responsive cis-Regulatory Sequences in Arabidopsis. *Plant Physiol.* **160**, 178-191 (2012).
262. Tian, F., Yang, D.C., Meng, Y.Q., Jin, J. & Gao, G. PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res* **48**, D1104-D1113 (2020).
263. Zhou, P. et al. Meta gene regulatory networks in maize highlight functionally relevant regulatory interactions. *Plant Cell* **32**, 1377-1396 (2020).
264. Teige, M. et al. The MKK2 pathway mediates cold and salt stress signaling in Arabidopsis. *Molecular Cell* **15**, 141-152 (2004).
265. Butelli, E. et al. Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nature Biotechnology* **26**, 1301-1308 (2008).
266. Nozawa, A. et al. Production and partial purification of membrane proteins using a liposome-supplemented wheat cell-free translation system. *BMC Biotechnol* **11**, 35 (2011).
267. Guseman, J.M., Webb, K., Srinivasan, C. & Dardick, C. DRO1 influences root system architecture in Arabidopsis and Prunus species. *Plant J* **89**, 1093-1105 (2017).
268. Zhang, Q. et al. Transcriptional regulatory networks controlling taste and aroma quality of apricot (Prunus armeniaca L.) fruit during ripening. *BMC Genomics* **20**, 1-15 (2019).
269. Guo, Z., Li, Y. & Ding, S.W. Small RNA-based antimicrobial immunity. *Nature Reviews Immunology* **19**, 31-44 (2019).
270. Peng, J. et al. 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* **400**, 256-261 (1999).
271. Bennett, A.B., Pankievicz, V.C.S. & Ané, J.M. A Model for Nitrogen Fixation in Cereal Crops. *Trends in Plant Science* **25**, 226-235 (2020).
272. Wang, J. et al. A single transcription factor promotes both yield and immunity in rice. *Science* **361**, 1026-1028 (2018).
273. Liu, M. et al. Inducible overexpression of Ideal Plant Architecture1 improves both yield and disease resistance in rice. *Nature Plants* **5**, 389-400 (2019).
274. Gao, Y. et al. Diversity and redundancy of the ripening regulatory networks revealed by the fruitENCODE and the new CRISPR/Cas9 CNR and NOR mutants. *Hortic Res* **6**, 39 (2019).
275. Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H. & Ohme-Takagi, M. Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. *Plant Cell* **13**, 1959-1968 (2001).
276. Hiratsu, K., Matsui, K., Koyama, T. & Ohme-Takagi, M. Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in Arabidopsis. *Plant J* **34**, 733-739 (2003).
277. Mahfouz, M.M. et al. Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. *Plant Molecular Biology* **78**, 311-321 (2012).
278. Cen, H. et al. Overexpression of a Chimeric Gene, OsDST-SRDX, Improved Salt Tolerance of Perennial Ryegrass. *Sci Rep* **6**, 27320 (2016).
279. Piatek, A. et al. RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant Biotechnol J* **13**, 578-589 (2015).
280. Matsui, K., Umemura, Y. & Ohme-Takagi, M. AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in Arabidopsis. *Plant J* **55**, 954-967 (2008).
281. Ikeda, M. & Ohme-Takagi, M. A novel group of transcriptional repressors in Arabidopsis. *Plant Cell Physiol* **50**, 970-975 (2009).
282. Campbell, M.E., Palfreyman, J.W. & Preston, C.M. Identification of herpes simplex virus DNA sequences which encode a trans-acting polypeptide responsible for stimulation of immediate early transcription. *J Mol Biol* **180**, 1-19 (1984).
283. Sadowski, I., Ma, J., Triezenberg, S. & Ptashne, M. GAL4-VP16 is an unusually potent transcriptional activator. *Nature* **335**, 563-564 (1988).

284. Liu, W. et al. Bacterial pathogen phytosensing in transgenic tobacco and Arabidopsis plants. *Plant Biotechnology Journal* **11**, 43-52 (2013).
285. Weinmann, P., Gossen, M., Hillen, W., Bujard, H. & Gatz, C. A chimeric transactivator allows tetracycline-responsive gene expression in whole plants. *The Plant Journal* **5**, 559-569 (1994).
286. Gardner, M.J. et al. GAL4 GFP enhancer trap lines for analysis of stomatal guard cell development and gene expression. *J Exp Bot* **60**, 213-226 (2009).
287. Zuo, J., Niu, Q.-W. & Chua, N.-H. An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *The Plant Journal* **24**, 265-273 (2000).
288. Gander, M.W., Vrana, J.D., Voje, W.E., Carothers, J.M. & Klavins, E. Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. *Nature Communications* **8** (2017).
289. Nielsen, A.A. & Voigt, C.A. Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol Syst Biol* **10**, 763 (2014).
290. Lowder, L.G. et al. A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. *Plant Physiol* **169**, 971-985 (2015).
291. Holmes-Davis, R. et al. Gene regulation in planta by plant-derived engineered zinc finger protein transcription factors. *Plant Molecular Biology* **57**, 411-423 (2005).
292. Van Eenennaam, A.L. et al. Elevation of seed alpha-tocopherol levels using plant-based transcription factors targeted to an endogenous locus. *Metabolic Engineering* **6**, 101-108 (2004).
293. Dandekar, A.M. et al. An engineered innate immune defense protects grapevines from Pierce disease. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 3721-3725 (2012).
294. Boch, J. & Bonas, U. Xanthomonas AvrBs3 Family-Type III Effectors: Discovery and Function. *Annual Review of Phytopathology* **48**, 419-436 (2010).
295. Raikhel, N. Nuclear targeting in plants. *Plant Physiol* **100**, 1627-1632 (1992).
296. Iacopino, S. et al. A Synthetic Oxygen Sensor for Plants Based on Animal Hypoxia Signaling. *Plant Physiol.* **179**, 986-1000 (2019).
297. Ellington, A.D. & Szostak, J.W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818-822 (1990).
298. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505-510 (1990).
299. Shanidze, N., Lenkeit, F., Hartig, J.S. & Funck, D. A Theophylline-Responsive Riboswitch Regulates Expression of Nuclear-Encoded Genes. *Plant Physiol.* **182**, 123-135 (2020).
300. Wu, M.F. et al. Auxin-regulated chromatin switch directs acquisition of flower primordium founder fate. *eLife* **4**, 1-20 (2015).
301. Nemhauser, J.L. & Torii, K.U. Plant synthetic biology for molecular engineering of signalling and development. *Nature Plants* **2**, 1-7 (2016).
302. Walia, A., Waadt, R. & Jones, A.M. Genetically Encoded Biosensors in Plants: Pathways to Discovery. *Annu Rev Plant Biol* **69**, 497-524 (2018).
303. Hammond, J.P. et al. Changes in gene expression in Arabidopsis shoots during phosphate starvation and the potential for developing smart plants. *Plant Physiol.* **132**, 578-596 (2003).
304. Kamiya, T., Yamagami, M., Hirai, M.Y. & Fujiwara, T. Establishment of an in planta magnesium monitoring system using CAX3 promoter-luciferase in Arabidopsis. *J Exp Bot* **63**, 355-363 (2012).
305. Krizek, B.A., Prost, V., Joshi, R.M., Stoming, T. & Glenn, T.C. Developing transgenic arabidopsis plants to be metal-specific bioindicators. *Environmental Toxicology and Chemistry* **22**, 175-181 (2003).
306. Zarka, D.G., Vogel, J.T., Cook, D. & Thomashow, M.F. Cold Induction of Arabidopsis CBF Genes Involves Multiple ICE (Inducer of CBF Expression) Promoter Elements and a Cold-Regulatory Circuit That Is Desensitized by Low Temperature. *Plant Physiol.* **133**, 910-918 (2003).
307. Brey, L.F. et al. Metabolic engineering of Synechocystis sp. PCC 6803 for the production of aromatic amino acids and derived phenylpropanoids. *Metabolic Engineering* **57**, 129-139 (2020).

308. Nishizawa, A. et al. Arabidopsis heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *Plant J.* **48**, 535-547 (2006).
309. Badran, A.H. et al. Continuous evolution of Bacillus thuringiensis toxins overcomes insect resistance. *Nature* **533**, 58-63 (2016).
310. Rushton, P.J., Reinstädler, A., Lipka, V., Lippok, B. & Somssich, I.E. Synthetic Plant Promoters Containing Defined Regulatory Elements Provide Novel Insights into Pathogen- and Wound-Induced Signaling. *The Plant Cell* **14**, 749-762 (2002).
311. Cheng, W. et al. Host-induced gene silencing of an essential chitin synthase gene confers durable resistance to Fusarium head blight and seedling blight in wheat. *Plant Biotechnology Journal* **13**, 1335-1345 (2015).
312. Higo, K., Ugawa, Y., Iwamoto, M. & Korenaga, T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research* **27**, 297-300 (1999).
313. Lescot, M. et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research* **30**, 325-327 (2002).
314. Hieno, A. et al. ppdb: plant promoter database version 3.0. *Nucleic Acids Research* **42**, D1188-D1192 (2014).
315. Chow, C.-N. et al. PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. *Nucleic Acids Research* **44**, D1154-D1160 (2016).
316. Wang, R. et al. Novel green tissue-specific synthetic promoters and cis-regulatory elements in rice. *Sci Rep-Uk* **5**, 1-11 (2015).
317. Bai, J. et al. Comprehensive construction strategy of bidirectional green tissue-specific synthetic promoters. *Plant Biotechnology Journal* **18**, 668-678 (2020).
318. Khakhar, A., Leydon, A.R., Lemmex, A.C., Klavins, E. & Nemhauser, J.L. Synthetic hormone-responsive transcription factors can monitor and re-program plant development. *eLife* **7** (2018).
319. Annaluru, N. et al. Total synthesis of a functional designer eukaryotic chromosome. *Science* **344**, 55-58 (2014).
320. Segall-Shapiro, T.H., Meyer, A.J., Ellington, A.D., Sontag, E.D. & Voigt, C.A. A 'resource allocator' for transcription based on a highly fragmented T7 RNA polymerase. *Mol Syst Biol* **10**, 742 (2014).
321. Pinto, F., Thornton, E.L. & Wang, B. An expanded library of orthogonal split inteins enables modular multi-peptide assemblies. *Nat Commun* **11**, 1529 (2020).
322. Müller, K. et al. A red light-controlled synthetic gene expression switch for plant systems. *Mol. BioSyst.* **10**, 1679-1688 (2014).
323. Cuthbertson, L. & Nodwell, J.R. The TetR Family of Regulators. *Microbiol. Mol. Biol. Rev.* **77**, 440-475 (2013).
324. Gatz, C., Frohberg, C. & Wendenburg, R. Stringent repression and homogeneous de-repression by tetracycline of a modified CaMV 35S promoter in intact transgenic tobacco plants. *Plant J* **2**, 397-404 (1992).
325. Arst, H.N., Jr., Holden, D.W. & Caddick, M.X. Evolution of transcription-regulating proteins: caveat lector! *Gene* **173 Suppl**, S1-4 (1997).
326. Persad, R. et al. The Q-System as a Synthetic Transcriptional Regulator in Plants. *Front Plant Sci* **11** (2020).
327. J, Z. & Nh, C. Chemical-inducible Systems for Regulated Expression of Plant Genes. *Current opinion in biotechnology* (2000).
328. Samalova, M., Brzobohaty, B. & Moore, I. pOp6/LhGR: a stringently regulated and highly responsive dexamethasone-inducible gene expression system for tobacco. *The Plant Journal* **41**, 919-935 (2005).
329. Padidam, M., Gore, M., Lily Lu, D. & Smirnova, O. Chemical-Inducible, Ecdysone Receptor-Based Gene Expression System for Plants. *Transgenic Research* **12**, 101-109 (2003).
330. Tavva, V.S., Dinkins, R.D., Palli, S.R. & Collins, G.B. Development of a methoxyfenozide-responsive gene switch for applications in plants. *Plant J* **45**, 457-469 (2006).

331. Martinez, A., Sparks, C., Hart, C.A., Thompson, J. & Jepson, I. Ecdysone agonist inducible transcription in transgenic tobacco plants. *Plant J* **19**, 97-106 (1999).
332. Omelina, E.S. et al. Optogenetic and Chemical Induction Systems for Regulation of Transgene Expression in Plants: Use in Basic and Applied Research. *Int J Mol Sci* **23** (2022).
333. Ce, T. et al. Computational Design of Ligand-Binding Proteins With High Affinity and Selectivity. *Nature* (2013).
334. Tang, S.Y. & Cirino, P.C. Design and application of a mevalonate-responsive regulatory protein. *Angewandte Chemie - International Edition* **50**, 1084-1086 (2011).
335. Bick, M.J. et al. Computational design of environmental sensors for the potent opioid fentanyl. *eLife* **6**, e28909 (2017).
336. Feng, J. et al. A general strategy to construct small molecule biosensors in eukaryotes. *eLife* **4**, e10606 (2015).
337. Tinberg, C.E. et al. Computational design of ligand-binding proteins with high affinity and selectivity. *Nature* **501**, 212-216 (2013).
338. Uchida, N. et al. Chemical hijacking of auxin signaling with an engineered auxin-TIR1 pair. *Nature Chemical Biology* **14**, 299-305 (2018).
339. Park, S.-Y. et al. Agrochemical control of plant water use using engineered abscisic acid receptors. *Nature* **520**, 545-548 (2015).
340. Antunes, M.S. et al. Engineering key components in a synthetic eukaryotic signal transduction pathway. *Molecular Systems Biology* **5**, 270 (2009).
341. Antunes, M.S. et al. Programmable Ligand Detection System in Plants through a Synthetic Signal Transduction Pathway. *PLoS ONE* **6** (2011).
342. Jang, S., Jang, S., Noh, M.H., Lim, H.G. & Jung, G.Y. Novel Hybrid Input Part Using Riboswitch and Transcriptional Repressor for Signal Inverting Amplifier. *ACS Synth Biol* **7**, 2199-2204 (2018).
343. Ramundo, S., Rahire, M., Schaad, O. & Rochaix, J.-D. Repression of Essential Chloroplast Genes Reveals New Signaling Pathways and Regulatory Feedback Loops in Chlamydomonas. *The Plant Cell* **25**, 167-186 (2013).
344. Bocobza, S. et al. Riboswitch-dependent gene regulation and its evolution in the plant kingdom. *Genes Dev* **21**, 2874-2879 (2007).
345. Emadpour, M., Karcher, D. & Bock, R. Boosting riboswitch efficiency by RNA amplification. *Nucleic Acids Res* **43**, e66 (2015).
346. Slusarczyk, A.L., Lin, A. & Weiss, R. Foundations for the design and implementation of synthetic genetic circuits. *Nat Rev Genet* **13**, 406-420 (2012).
347. Bittihn, P., Din, M.O., Tsimring, L.S. & Hasty, J. Rational engineering of synthetic microbial systems: from single cells to consortia. *Curr Opin Microbiol* **45**, 92-99 (2018).
348. Xie, M. & Fussenegger, M. Designing cell function: assembly of synthetic gene circuits for cell biology applications. *Nat Rev Mol Cell Biol* **19**, 507-525 (2018).
349. Ceroni, F. & Ellis, T. The challenges facing synthetic biology in eukaryotes. *Nat Rev Mol Cell Biol* **19**, 481-482 (2018).
350. Brophy, J.A.N., LaRue, T. & Dinneny, J.R. Understanding and engineering plant form. *Seminars in Cell & Developmental Biology* **79**, 68-77 (2018).
351. Lowder, L.G. et al. Robust Transcriptional Activation in Plants Using Multiplexed CRISPR-Act2.0 and mTALE-Act Systems. *Molecular Plant* **11**, 245-256 (2018).
352. Reis, R.S., Litholdo, C.G., Bally, J., Roberts, T.H. & Waterhouse, P.M. A conditional silencing suppression system for transient expression. *Sci Rep-Uk* **8**, 1-7 (2018).
353. Cox, R.S., Surette, M.G. & Elowitz, M.B. Programming gene expression with combinatorial promoters. *Molecular Systems Biology* **3** (2007).
354. Brophy, J.A., Magallon, K.J., Kniazev, K. & Dinneny, J.R. Synthetic genetic circuits enable reprogramming of plant roots. *bioRxiv* (2022).
355. Lloyd, J.P.B. et al. Synthetic memory circuits for stable cell reprogramming in plants. *Nat Biotechnol* (2022).

356. Hart, Y. & Alon, U. The utility of paradoxical components in biological circuits. *Mol Cell* **49**, 213-221 (2013).
357. Hart, Y., Mayo, A.E., Milo, R. & Alon, U. Robust control of PEP formation rate in the carbon fixation pathway of C(4) plants by a bi-functional enzyme. *BMC Syst Biol* **5**, 171 (2011).
358. Rodrigo, G., Bajic, D., Elola, I. & Poyatos, J.F. Antagonistic autoregulation speeds up a homogeneous response in Escherichia coli. *Sci Rep* **6**, 36196 (2016).
359. Schreiber, T., Prange, A., Hoppe, T. & Tissier, A. Split-tale: A tale-based two-component system for synthetic biology applications in planta. *Plant Physiol.* **179**, 1001-1012 (2019).
360. Cordero, T. et al. Boolean Computation in Plants Using Post-translational Genetic Control and a Visual Output Signal. *ACS Synth Biol* **7**, 2322-2330 (2018).
361. Purnick, P.E.M. & Weiss, R. The second wave of synthetic biology: From modules to systems. *Nature Reviews Molecular Cell Biology* **10**, 410-422 (2009).
362. Medford, J.I. & Prasad, A. Towards programmable plant genetic circuits. *The Plant Journal* **87**, 139-148 (2016).
363. Kassaw, T.K., Donayre-Torres, A.J., Antunes, M.S., Morey, K.J. & Medford, J.I. Engineering synthetic regulatory circuits in plants. *Plant Science* **273**, 13-22 (2018).
364. Siuti, P., Yazbek, J. & Lu, T.K. Synthetic circuits integrating logic and memory in living cells. *Nature Biotechnology* **31**, 448-452 (2013).
365. Yang, L. et al. Permanent genetic memory with >1-byte capacity. *Nature Methods* **11**, 1261-1266 (2014).
366. Bernabé-Orts, J.M. et al. A memory switch for plant synthetic biology based on the phage ϕC31 integration system. *Nucleic Acids Research* **48**, 3379-3394 (2020).
367. Sun, B. et al. Timing mechanism dependent on cell division is invoked by Polycomb eviction in plant stem cells. *Science* **343**, 1248559 (2014).
368. Feng, Y. et al. A Three-Component Gene Expression System and Its Application for Inducible Flavonoid Overproduction in Transgenic Arabidopsis thaliana. *PLoS ONE* **6** (2011).
369. Okada, R., Nemoto, Y., Endo-Higashi, N. & Izawa, T. Synthetic control of flowering in rice independent of the cultivation environment. *Nature Plants* **3**, 1-7 (2017).
370. Kazan, K. & Lyons, R. The link between flowering time and stress tolerance. *J Exp Bot* **67**, 47-60 (2016).
371. Jung, C. & Muller, A.E. Flowering time control and applications in plant breeding. *Trends Plant Sci* **14**, 563-573 (2009).
372. Liang, Y.-C., Reid, M.S. & Jiang, C.-Z. Controlling plant architecture by manipulation of gibberellic acid signalling in petunia. *Horticulture Research* **1**, 1-6 (2014).
373. Mahfouz, M.M. et al. Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. *Plant Molecular Biology* **78**, 311-321 (2012).
374. Cong, L. et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819-823 (2013).
375. Li, Z. et al. A potent Cas9-derived gene activator for plant and mammalian cells. *Nat Plants* **3**, 930-936 (2017).
376. Gupta, M. et al. Transcriptional activation of Brassica napus beta-ketoacyl-ACP synthase II with an engineered zinc finger protein transcription factor. *Plant Biotechnol J* **10**, 783-791 (2012).
377. Lowder, L.G. et al. Robust Transcriptional Activation in Plants Using Multiplexed CRISPR-Act2.0 and mTALE-Act Systems. *Mol Plant* **11**, 245-256 (2018).
378. Guo, Q., Liu, Q., Smith, N.A., Liang, G. & Wang, M.B. RNA Silencing in Plants: Mechanisms, Technologies and Applications in Horticultural Crops. *Curr Genomics* **17**, 476-489 (2016).
379. Ogita, S., Uefuji, H., Yamaguchi, Y., Koizumi, N. & Sano, H. Producing decaffeinated coffee plants. *Nature* **423**, 823 (2003).
380. Waterhouse, P.M., Graham, H.W. & Wang, M.B. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 13959-13964 (1998).

381. Ledger, S.E., Janssen, B.J., Karunairetnam, S., Wang, T.C. & Snowden, K.C. Modified CAROTENOID CLEAVAGE DIOXYGENASE8 expression correlates with altered branching in kiwifruit (Actinidia chinensis). *New Phytol* **188**, 803-813 (2010).
382. Liu, S. & Yoder, J.I. Chemical induction of hairpin RNAi molecules to silence vital genes in plant roots. *Sci Rep-Uk* **6**, 37711 (2016).
383. Nakatsuka, T. et al. Genetic engineering of novel flower colour by suppression of anthocyanin modification genes in gentian. *J Plant Physiol* **167**, 231-237 (2010).
384. Zhao, X. et al. Global identification of Arabidopsis lncRNAs reveals the regulation of MAF4 by a natural antisense RNA. *Nature Communications* **9**, 1-12 (2018).
385. Barros, J., Temple, S. & Dixon, R.A. Development and commercialization of reduced lignin alfalfa. *Curr Opin Biotechnol* **56**, 48-54 (2019).
386. Guo, D., Chen, F., Inoue, K., Blount, J.W. & Dixon, R.A. Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa. impacts on lignin structure and implications for the biosynthesis of G and S lignin. *Plant Cell* **13**, 73-88 (2001).
387. Abbott, J.C. et al. Simultaneous suppression of multiple genes by single transgenes. Down-regulation of three unrelated lignin biosynthetic genes in tobacco. *Plant Physiol.* **128**, 844-853 (2002).
388. Agrawal, R.K., Patel, R.K., shah, V., Nainiwal, L. & Trivedi, B. Hydroxyurea in Sickle Cell Disease: Drug Review. *Indian J Hematol Blood Transfus* **30**, 91-96 (2014).
389. Manavski, N. et al. In vivo stabilization of endogenous chloroplast RNAs by customized artificial pentatricopeptide repeat proteins. *BioRxiv*, 10.1101/2020.1111.1121.39274-39210.31101/32020.39211.39221.39274 (2020).
390. Aman, R. et al. RNA virus interference via CRISPR/Cas13a system in plants. *Genome Biol* **19**, 1 (2018).
391. Abudayyeh, O.O. et al. RNA targeting with CRISPR-Cas13. *Nature* **550**, 280-284 (2017).
392. Mahas, A., Neal Stewart, C. & Mahfouz, M.M. Harnessing CRISPR/Cas systems for programmable transcriptional and post-transcriptional regulation. *Biotechnology Advances* **36**, 295-310 (2018).
393. Hickey, S.F. et al. Transgene regulation in plants by alternative splicing of a suicide exon. *Nucleic Acids Res* **40**, 4701-4710 (2012).
394. Gonzalez, T.L. et al. Tight regulation of plant immune responses by combining promoter and suicide exon elements. *Nucleic acids research* **43**, 7152-7161 (2015).
395. Johnson, L.M. et al. SRA-and SET-domain-containing proteins link RNA polymerase v occupancy to DNA methylation. *Nature* **507**, 124-128 (2014).
396. Yang, X. et al. Plant Biosystems Design Research Roadmap 1.0. *BioDesign Research* **2020**, 1-38 (2020).
397. Navarro, F.J. & Baulcombe, D.C. MiRNA-Mediated Regulation of Synthetic Gene Circuits in the Green Alga Chlamydomonas reinhardtii. *ACS Synth Biol* (2019).
398. Pfeiffer, A. et al. A cell-free system for light-dependent nuclear import of phytochrome. *Plant Journal* **57**, 680-689 (2009).
399. Gordon, G. & Pfleger, B. Regulatory tools for controlling gene expression in cyanobacteria. *Adv Exp Med Biol* **1080**, 281-315 (2018).
400. Schaumberg, K.A. et al. Quantitative characterization of genetic parts and circuits for plant synthetic biology. *Nature Methods* **13**, 94-100 (2016).
401. Boehm, C.R., Pollak, B., Purswani, N., Patron, N. & Haseloff, J. Synthetic Botany. *Cold Spring Harb Perspect Biol* **9** (2017).
402. Kong, F., Yamaoka, Y., Ohama, T., Lee, Y. & Li-Beisson, Y. Molecular genetic tools and emerging synthetic biology strategies to increase cellular oil content in chlamydomonas reinhardtii. *Plant and Cell Physiology* **60**, 1184-1196 (2019).
403. Draper, J. et al. Brachypodium distachyon. A new model system for functional genomics in grasses. *Plant Physiol* **127**, 1539-1555 (2001).

404. Huang, P., Shyu, C., Coelho, C.P., Cao, Y. & Brutnell, T.P. Setaria viridis as a Model System to Advance Millet Genetics and Genomics. *Front Plant Sci* **7**, 1781 (2016).
405. Delmans, M., Pollak, B. & Haseloff, J. MarpoDB: An open registry for Marchantia polymorpha genetic parts. *Plant and Cell Physiology* **58**, e5-e5 (2017).
406. Sauret-Gueto, S. et al. Systematic Tools for Reprogramming Plant Gene Expression in a Simple Model, Marchantia polymorpha. *ACS Synth Biol* **9**, 864-882 (2020).
407. Dlugosz, E.M., Lenaghan, S.C. & Stewart, C.N. A Robotic Platform for High-throughput Protoplast Isolation and Transformation. *Journal of Visualized Experiments* (2016).
408. Lenaghan, S.C. & Neal Stewart, C. An automated protoplast transformation system. *Methods in Molecular Biology* **1917**, 355-363 (2019).
409. Yu, Z. et al. Droplet-based microfluidic analysis and screening of single plant cells. *PLoS One* **13**, e0196810 (2018).
410. Carqueijeiro, I. et al. Beyond the semi-synthetic artemisinin: metabolic engineering of plant-derived anti-cancer drugs. *Current Opinion in Biotechnology* **65**, 17-24 (2020).
411. Del Valle, I. et al. Soil organic matter attenuates the efficacy of flavonoid-based plant-microbe communication. *Science Advances* **6**, eaax8254-eaax8254 (2020).
412. Sawasaki, T. et al. A bilayer cell-free protein synthesis system for high-throughput screening of gene products. *FEBS Lett* **514**, 102-105 (2002).
413. Fogeron, M.L., Lecoq, L., Cole, L., Harbers, M. & Bockmann, A. Easy Synthesis of Complex Biomolecular Assemblies: Wheat Germ Cell-Free Protein Expression in Structural Biology. *Front Mol Biosci* **8** (2021).
414. Dudley, Q.M. et al. Biofoundry-assisted expression and characterisation of plant proteins. *bioRxiv*, 2021.2003.2011.434954 (2021).
415. Buntru, M., Vogel, S., Spiegel, H. & Schillberg, S. Tobacco BY-2 cell-free lysate: an alternative and highly-productive plant-based in vitro translation system. *BMC Biotechnol* **14**, 37 (2014).
416. Buntru, M., Vogel, S., Stoff, K., Spiegel, H. & Schillberg, S. A versatile coupled cell-free transcription-translation system based on tobacco BY-2 cell lysates. *Biotechnol Bioeng* **112**, 867-878 (2015).
417. Hirose, T. & Sugiura, M. Cis-acting elements and trans-acting factors for accurate translation of chloroplast psbA mRNAs: Development of an in vitro translation system from tobacco chloroplasts. *Embo J* **15**, 1687-1695 (1996).
418. Yukawa, M., Kuroda, H. & Sugiura, M. A new in vitro translation system for non-radioactive assay from tobacco chloroplasts: effect of pre-mRNA processing on translation in vitro. *Plant J* **49**, 367-376 (2007).
419. Gruissem, W. & Zurawski, G. Analysis of Promoter Regions for the Spinach Chloroplast Rbcl, Atpb and Psba Genes. *Embo J* **4**, 3375-3383 (1985).
420. Hanleybowdoin, L. & Chua, N.H. Chloroplast Promoters. *Trends in Biochemical Sciences* **12**, 67-70 (1987).
421. De Paepe, B., Maertens, J., Vanholme, B. & De Mey, M. Chimeric LysR-Type Transcriptional Biosensors for Customizing Ligand Specificity Profiles toward Flavonoids. *ACS Synth Biol* **8**, 318-331 (2019).
422. You, Y.-S. et al. Use of Bacterial Quorum-Sensing Components to Regulate Gene Expression in Plants. *Plant Physiol.* **140**, 1205-1212 (2006).
423. Khakhar, A., Leydon, A.R., Lemmex, A.C., Klavins, E. & Nemhauser, J.L. Synthetic hormone-responsive transcription factors can monitor and re-program plant development. *eLife* **7**.
424. Bais, H.P., Weir, T.L., Perry, L.G., Gilroy, S. & Vivanco, J.M. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol* **57**, 233-266 (2006).
425. Carvalhais, L.C. et al. Root exudation of sugars, amino acids, and organic acids by maize as affected by nitrogen, phosphorus, potassium, and iron deficiency. *Journal of Plant Nutrition and Soil Science* **174**, 3-11 (2011).

426. Harbort, C.J. et al. Root-Secreted Coumarins and the Microbiota Interact to Improve Iron Nutrition in Arabidopsis. *Cell Host & Microbe* **28**, 825-+ (2020).
427. Miller, W.G., Brandl, M.T., Quinones, B. & Lindow, S.E. Biological sensor for sucrose availability: Relative sensitivities of various reporter genes. *Appl Environ Microbiol* **67**, 1308-1317 (2001).
428. Leveau, J.H. & Lindow, S.E. Appetite of an epiphyte: quantitative monitoring of bacterial sugar consumption in the phyllosphere. *Proc Natl Acad Sci U S A* **98**, 3446-3453 (2001).
429. Jaeger, C.H., 3rd, Lindow, S.E., Miller, W., Clark, E. & Firestone, M.K. Mapping of sugar and amino acid availability in soil around roots with bacterial sensors of sucrose and tryptophan. *Appl Environ Microbiol* **65**, 2685-2690 (1999).
430. Bringhurst, R.M., Cardon, Z.G. & Gage, D.J. Galactosides in the rhizosphere: Utilization by sinorhizobium meliloti and development of a biosensor. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 4540-4545 (2001).
431. Tonti-Filippini, J., Nevill, P.G., Dixon, K. & Small, I. What can we do with 1000 plastid genomes? *Plant J* **90**, 808-818 (2017).
432. Pini, F. et al. Bacterial Biosensors for in Vivo Spatiotemporal Mapping of Root Secretion. *Plant Physiol.* **174**, 1289-1306 (2017).
433. Jang, S. et al. Development of Artificial Riboswitches for Monitoring of Naringenin In Vivo. *ACS Synth Biol* **6**, 2077-2085 (2017).
434. Ryu, M.-H. et al. Control of nitrogen fixation in bacteria that associate with cereals. *Nature Microbiology* **5**, 314-330 (2020).
435. Meyer, A.J., Segall-Shapiro, T.H., Glassey, E., Zhang, J. & Voigt, C.A. Escherichia coli "Marionette" strains with 12 highly optimized small-molecule sensors. *Nat Chem Biol* **15**, 196-204 (2019).
436. Horbal, L., Fedorenko, V. & Luzhetskyy, A. Novel and tightly regulated resorcinol and cumate-inducible expression systems for Streptomyces and other actinobacteria. *Appl Microbiol Biotechnol* **98**, 8641-8655 (2014).
437. Hanko, E.K.R. et al. A genome-wide approach for identification and characterisation of metabolite-inducible systems. *Nature Communications* **11** (2020).
438. Siedler, S. et al. Development of a Bacterial Biosensor for Rapid Screening of Yeast p-Coumaric Acid Production. *ACS Synth Biol* **6**, 1860-1869 (2017).
439. Jiang, T., Li, C.Y. & Yan, Y.J. Optimization of a p-Coumaric Acid Biosensor System for Versatile Dynamic Performance. *ACS Synth Biol* **10**, 132-144 (2021).
440. Machado, L.F.M., Currin, A. & Dixon, N. Directed evolution of the PcaV allosteric transcription factor to generate a biosensor for aromatic aldehydes. *Journal of Biological Engineering* **13** (2019).
441. Dietrich, J.A., Shis, D.L., Alikhani, A. & Keasling, J.D. Transcription Factor-Based Screens and Synthetic Selections for Microbial Small-Molecule Biosynthesis. *ACS Synth Biol* **2**, 47-58 (2013).
442. Ganesh, M., Hemalatha, P., Peng, M.M., Cha, W.S. & Jang, H.T. Zr-Fumarate MOF a Novel CO2-Adsorbing Material: Synthesis and Characterization. *Aerosol Air Qual Res* **14**, 1605-1612 (2014).
443. Ganesh, I., Ravikumar, S., Lee, S.H., Park, S.J. & Hong, S.H. Engineered fumarate sensing Escherichia coli based on novel chimeric two-component system. *Journal of Biotechnology* **168**, 560-566 (2013).
444. Thompson, M.G. et al. Robust Characterization of Two Distinct Glutarate Sensing Transcription Factors of Pseudomonas putida L-Lysine Metabolism. *ACS Synth Biol* **8**, 2385-2396 (2019).
445. Gonzalez, L.M., Mukhitov, N. & Voigt, C.A. Resilient living materials built by printing bacterial spores. *Nature Chemical Biology* **16**, 126-+ (2020).
446. Dimas, R.P., Jiang, X.L., de la Paz, J.A., Morcos, F. & Chan, C.T.Y. Engineering repressors with coevolutionary cues facilitates toggle switches with a master reset. *Nucleic Acids Research* **47**, 5449-5463 (2019).

447. Chew, Y.H. et al. Multiscale digital Arabidopsis predicts individual organ and whole-organism growth (vol 111, pg E4127, 2014). *Proceedings of the National Academy of Sciences of the United States of America* **112**, E2556-E2556 (2015).
448. Rogers, J.K. et al. Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Research* **43**, 7648-7660 (2015).
449. Siedler, S., Stahlhut, S.G., Malla, S., Maury, J. & Neves, A.R. Novel biosensors based on flavonoid-responsive transcriptional regulators introduced into Escherichia coli. *Metabolic Engineering* **21**, 2-8 (2014).
450. Binder, S. et al. A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. *Genome Biol* **13** (2012).
451. Ellefson, J.W., Ledbetter, M.P. & Ellington, A.D. Directed evolution of a synthetic phylogeny of programmable Trp repressors. *Nature Chemical Biology* **14**, 361-+ (2018).
452. Landry, B.P., Palanki, R., Dyulgyarov, N., Hartsough, L.A. & Tabor, J.J. Phosphatase activity tunes two-component system sensor detection threshold. *Nature Communications* **9**, 1-10 (2018).
453. Ravikumar, S., David, Y., Park, S.J. & Choi, J.I. A Chimeric Two-Component Regulatory System-Based Escherichia coli Biosensor Engineered to Detect Glutamate. *Appl Biochem Biotech* **186**, 335-349 (2018).
454. Sun, H.H., Zhao, H.M. & Ang, E.L. A New Biosensor for Stilbenes and a Cannabinoid Enabled by Genome Mining of a Transcriptional Regulator. *ACS Synth Biol* **9**, 698-705 (2020).
455. McClune, C.J., Alvarez-Buylla, A., Voigt, C.A. & Laub, M.T. Engineering orthogonal signalling pathways reveals the sparse occupancy of sequence space. *Nature* **574**, 702-706 (2019).
456. Pandey, R.P., Parajuli, P., Koffas, M.A.G. & Sohng, J.K. Microbial production of natural and non-natural flavonoids: Pathway engineering, directed evolution and systems/synthetic biology. *Biotechnology Advances* **34**, 634-662 (2016).
457. Smee, M.R., Real-Ramirez, I. & Hendry, T.A. Insects as phyllosphere microbiome engineers: effects of aphids on a plant pathogen. *bioRxiv*, 797738 (2019).
458. Frago, E. et al. Symbionts protect aphids from parasitic wasps by attenuating herbivore-induced plant volatiles. *Nature Communications* **8**, 1-9 (2017).
459. Bruce, T.J.A. et al. The first crop plant genetically engineered to release an insect pheromone for defence. *Sci Rep-Uk* **5**, 1-9 (2015).
460. Ding, P.T. & Ding, Y.L. Stories of Salicylic Acid: A Plant Defense Hormone. *Trends in Plant Science* **25**, 549-565 (2020).
461. Bloch, S.E. et al. Biological nitrogen fixation in maize: optimizing nitrogenase expression in a root-associated diazotroph. *J Exp Bot* (2020).
462. McGenity, T.J., Crombie, A.T. & Murrell, J.C. Microbial cycling of isoprene, the most abundantly produced biological volatile organic compound on Earth. *Isme Journal* **12**, 931-941 (2018).
463. Fini, A. et al. Isoprene Responses and Functions in Plants Challenged by Environmental Pressures Associated to Climate Change. *Front Plant Sci* **8** (2017).
464. Kim, S.K. et al. A Genetically Encoded Biosensor for Monitoring Isoprene Production in Engineered Escherichia coli. *ACS Synth Biol* **7**, 2379-2390 (2018).
465. Wang, M. et al. Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection. *Nature Plants* **2**, 1-10 (2016).
466. Weiberg, A. et al. Fungal Small RNAs Suppress Plant Immunity by Hijacking Host RNA Interference Pathways. *Science* **342**, 118-123 (2013).
467. Koch, A. et al. An RNAi-Based Control of Fusarium graminearum Infections Through Spraying of Long dsRNAs Involves a Plant Passage and Is Controlled by the Fungal Silencing Machinery. *PLOS Pathogens* **12**, e1005901 (2016).
468. Zhang, T. et al. Cotton plants export microRNAs to inhibit virulence gene expression in a fungal pathogen. *Nature Plants* **2**, 1-6 (2016).

469. Abdel-Ghany, S.E., Day, I., Heuberger, A.L., Broeckling, C.D. & Reddy, A.S. Production of Phloroglucinol, a Platform Chemical, in Arabidopsis using a Bacterial Gene. *Sci Rep* **6**, 38483 (2016).
470. Meyer, A. et al. Organism Engineering for the Bioproduction of the Triaminotrinitrobenzene (TATB) Precursor Phloroglucinol (PG). *ACS Synth Biol* **8**, 2746-2755 (2019).
471. Geddes, B.A. et al. Engineering transkingdom signalling in plants to control gene expression in rhizosphere bacteria. *Nature Communications* **10**, 1-11 (2019).
472. Zhalnina, K. et al. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat Microbiol* **3**, 470-480 (2018).
473. Ali, B., Sabri, A.N., Ljung, K. & Hasnain, S. Auxin production by plant associated bacteria: impact on endogenous IAA content and growth of Triticum aestivum L. *Lett Appl Microbiol* **48**, 542-547 (2009).
474. Fernandez-Llamosas, H. et al. Enhancing the Rice Seedlings Growth Promotion Abilities ofAzoarcussp. CIB by Heterologous Expression of ACC Deaminase to Improve Performance of Plants Exposed to Cadmium Stress. *Microorganisms* **8** (2020).
475. Trivedi, P.a.D.-B.M.a.T.C.a.H.K.a.A.I.C.a.S.B.K. Keystone microbial taxa regulate the invasion of a fungal pathogen in agro-ecosystems. *Soil Biology and Biochemistry* **111**, 10--14 (2017).
476. Alori, E.T.a.G.B.R.a.B.O.O. Microbial phosphorus solubilization and its potential for use in sustainable agriculture. *Front. Microbiol.* **8**, 1--8 (2017).
477. Lang, J. et al. Fitness costs restrict niche expansion by generalist niche-constructing pathogens. *The ISME Journal* **11**, 374-385 (2017).
478. Oger, P., Petit, A. & Dessaux, Y. Genetically engineered plants producing opines alter their biological environment. *Nature biotechnology* **15**, 369 (1997).
479. Savka, M.A. & Farrand, S.K. Modification of rhizobacterial populations by engineering bacterium utilization of a novel plant-produced resource. *Nat Biotechnol* **15**, 363-368 (1997).
480. Harlan, J.R. Crops and man. (American Society of Agronomy, 1992).
481. Simmonds, J. Community matters: A history of biological nitrogen fixation and nodulation research, 1965 to 1995. (Rensselaer Polytechnic Institute, 2007).
482. Erisman, J., Bleeker, A., Galloway, J. & Sutton, M. Reduced nitrogen in ecology and the environment. *Environmental pollution* **150**, 140-149 (2007).
483. Geddes, B.A. et al. Use of plant colonizing bacteria as chassis for transfer of N 2-fixation to cereals. *Current opinion in biotechnology* **32**, 216-222 (2015).
484. Zaim, S., Bekkar, A.A. & Belabid, L. in Rhizobium Biology and Biotechnology 25-37 (Springer, 2017).
485. Nadarajah, K.K. in Rhizobium Biology and Biotechnology 83-103 (Springer, 2017).
486. Gutierrez-Zamora, M. & Martınez-Romero, E. Natural endophytic association between Rhizobium etli and maize (Zea mays L.). *Journal of Biotechnology* **91**, 117-126 (2001).
487. McInroy, J.A. & Kloepper, J.W. Survey of indigenous bacterial endophytes from cotton and sweet corn. *Plant Soil* **173**, 337-342 (1995).
488. Ramachandran, V.K., East, A.K., Karunakaran, R., Downie, J.A. & Poole, P.S. Adaptation of Rhizobium leguminosarum to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. *Genome Biol* **12**, R106 (2011).
489. Frans, J. et al. in Nitrogen Fixation 33-44 (Springer, 1990).
490. Delmotte, N. et al. An integrated proteomics and transcriptomics reference data set provides new insights into the Bradyrhizobium japonicum bacteroid metabolism in soybean root nodules. *Proteomics* **10**, 1391-1400 (2010).
491. James, E. Nitrogen fixation in endophytic and associative symbiosis. *Field Crop Res* **65**, 197-209 (2000).
492. Ars‘ene, F., Katupitiya, S., Kennedy, I.R. & Elmerich, C. Use of IacZ Fusions to Study the Expression of nif Genes of Azospirillum brasilense. (1994).

493. Boddey, R.M. et al. in Management of Biological Nitrogen Fixation for the Development of More Productive and Sustainable Agricultural Systems 195-209 (Springer, 1995).
494. Bremer, E., Janzen, H. & Gilbertson, C. Evidence against associative N 2 fixation as a significant N source in long-term wheat plots. *Plant Soil* **175**, 13-19 (1995).
495. Chalk, P. The contribution of associative and symbiotic nitrogen fixation to the nitrogen nutrition of non-legumes. *Plant Soil* **132**, 29-39 (1991).
496. Hurek, T., Reinhold-Hurek, B., Van Montagu, M. & Kellenberger, E. Root colonization and systemic spreading of Azoarcus sp. strain BH72 in grasses. *Journal of bacteriology* **176**, 1913-1923 (1994).
497. James, E., Olivares, F., Baldani, J. & Döbereiner, J. Herbaspirillum, an endophytic diazotroph colonizing vascular tissue 3 Sorghum bicolor L. Moench. *J Exp Bot* **48**, 785-798 (1997).
498. Kapulnik, Y., Feldman, M., Okon, Y. & Henis, Y. Contribution of nitrogen fixed by Azospirillum to the N nutrition of spring wheat in Israel. *Soil Biology and Biochemistry* **17**, 509-515 (1985).
499. Katupitiya, S. et al. A mutant of Azospirillum brasilense Sp7 impaired in flocculation with a modified colonization pattern and superior nitrogen fixation in association with wheat. *Appl Environ Microbiol* **61**, 1987-1995 (1995).
500. Malik, K. et al. Association of nitrogen-fixing, plant-growth-promoting rhizobacteria (PGPR) with kallar grass and rice. *Plant Soil* **194**, 37-44 (1997).
501. Reinhold-Hurek, B. & Hurek, T. Interactions of gramineous plants with Azoarcus spp. and other diazotrophs: identification, localization, and perspectives to study their function. *Critical Reviews in Plant Sciences* **17**, 29-54 (1998).
502. Stoltzfus, J., So, R., Malarvithi, P., Ladha, J. & De Bruijn, F. Isolation of endophytic bacteria from rice and assessment of their potential for supplying rice with biologically fixed nitrogen. *Plant Soil* **194**, 25-36 (1997).
503. Triplett, E.W. Diazotrophic endophytes: progress and prospects for nitrogen fixation in monocots. *Plant Soil* **186**, 29-38 (1996).
504. Broek, A.V., Michiels, J., Van Gool, A. & Vanderleyden, J. Spatial-temporal colonization patterns of Azospirillum brasilense on the wheat root surface and expression of the bacterial nifH gene during association. *Mol. Plant-Microbe Interact.* **6**, 592-600 (1993).
505. Yanni, Y.G. et al. in Opportunities for Biological Nitrogen Fixation in Rice and Other Non-Legumes 99-114 (Springer, 1997).
506. Okon, Y. Azospirillum as a potential inoculant for agriculture. *Trends in Biotechnology* **3**, 223-228 (1985).
507. Cocking, E.C. Endophytic colonization of plant roots by nitrogen-fixing bacteria. *Plant Soil* **252**, 169-175 (2003).
508. Woodard, H. & Bly, A. Maize growth and yield responses to seed-inoculated N2-fixing bacteria under Dryland production conditions. *Journal of plant nutrition* **23**, 55-65 (2000).
509. Igiehon, N.O. & Babalola, O.O. Rhizosphere Microbiome Modulators: Contributions of Nitrogen Fixing Bacteria towards Sustainable Agriculture. *International journal of environmental research and public health* **15**, 574 (2018).
510. Biswas, J., Ladha, J. & Dazzo, F. Rhizobia inoculation improves nutrient uptake and growth of lowland rice. (2000).
511. Biswas, J.C., Ladha, J.K., Dazzo, F.B., Yanni, Y.G. & Rolfe, B.G. Rhizobial inoculation influences seedling vigor and yield of rice. *Agronomy Journal* **92**, 880-886 (2000).
512. Cummings, S.P. et al. Nodulation of Sesbania species by Rhizobium (Agrobacterium) strain IRBG74 and other rhizobia. *Environ. Microbiol.* **11**, 2510-2525 (2009).
513. Masson-Boivin, C., Giraud, E., Perret, X. & Batut, J. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends in Microbiology* **17**, 458-466 (2009).
514. Hoover, T.R., Imperial, J., Ludden, P.W. & Shah, V.K. Homocitrate is a component of the iron-molybdenum cofactor of nitrogenase. *Biochemistry-Us* **28**, 2768-2771 (1989).

515. Mandon, K. et al. Role of the fixGHI region of Azorhizobium caulinodans in free-living and symbiotic nitrogen fixation. **114**, 185-189 (1993).
516. Tsukada, S. et al. Comparative genome-wide transcriptional profiling of Azorhizobium caulinodans ORS571 grown under free-living and symbiotic conditions. **75**, 5037-5046 (2009).
517. Gopalaswamy, G., Kannaiyan, S., O'Callaghan, K.J., Davey, M.R. & Cocking, E.C. The xylem of rice (Oryza sativa) is colonized by Azorhizobium caulinodans. *Proc Biol Sci* **267**, 103-107 (2000).
518. Webster, G. et al. The flavonoid naringenin stimulates the intercellular colonization of wheat roots by Azorhizobium caulinodans. *Plant, Cell & Environment* **21**, 373-383 (1998).
519. Webster, G. et al. in Opportunities for Biological Nitrogen Fixation in Rice and Other Non-Legumes 115-122 (Springer, 1997).
520. Triplett, E.W., Kaeppler, S.M. & Chelius, M.K. (Google Patents, 2008).
521. Fujii, T. et al. Effect of inoculation withKlebsiella oxytoca andEnterobacter cloacae on dinitrogen fixation by rice-bacteria associations. *Plant Soil* **103**, 221-226 (1987).
522. El-Khawas, H. & Adachi, K. Identification and quantification of auxins in culture media of Azospirillum and Klebsiella and their effect on rice roots. *Biol Fertil Soils* **28**, 377-381 (1999).
523. Palus, J.A., Borneman, J., Ludden, P.W., Triplett, E.W.J.P. & soil A diazotrophic bacterial endophyte isolated from stems of Zea mays L. and Zea luxurians Iltis and Doebley. **186**, 135-142 (1996).
524. Rediers, H. et al. Development and application of a dapB-based in vivo expression technology system to study colonization of rice by the endophytic nitrogen-fixing bacterium Pseudomonas stutzeri A15. *Appl Environ Microbiol* **69**, 6864-6874 (2003).
525. Haahtela, K. & Korhonen, T.K. In vitro adhesion of N2-fixing enteric bacteria to roots of grasses and cereals. *Appl Environ Microbiol* **49**, 1186-1190 (1985).
526. Desnoues, N. et al. Nitrogen fixation genetics and regulation in a Pseudomonas stutzeri strain associated with rice. *Microbiology* **149**, 2251-2262 (2003).
527. Paulsen, I.T. et al. Complete genome sequence of the plant commensal Pseudomonas fluorescens Pf-5. *Nature biotechnology* **23**, 873 (2005).
528. Walsh, U.F., Morrissey, J.P. & O'Gara, F. Pseudomonas for biocontrol of phytopathogens: from functional genomics to commercial exploitation. *Current Opinion in Biotechnology* **12**, 289-295 (2001).
529. Dos Santos, P.C., Fang, Z., Mason, S.W., Setubal, J.C. & Dixon, R. Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC genomics* **13**, 162 (2012).
530. Boyd, E.S., Costas, A.M.G., Hamilton, T.L., Mus, F. & Peters, J.W. Evolution of molybdenum nitrogenase during the transition from anaerobic to aerobic metabolism. *Journal of bacteriology* **197**, 1690-1699 (2015).
531. Raymond, J., Siefert, J.L., Staples, C.R. & Blankenship, R.E. The natural history of nitrogen fixation. *Molecular biology and evolution* **21**, 541-554 (2004).
532. Rubio, L.M. & Ludden, P.W. Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Annu. Rev. Microbiol.* **62**, 93-111 (2008).
533. Iki, T., Aono, T. & Oyaizu, H. Evidence for functional differentiation of duplicated nifH genes in Azorhizobium caulinodans. *FEMS Microbiol Lett* **274**, 173-179 (2007).
534. Ratet, P., Pawlowski, K., Schell, J. & Bruijn, F. The Azorhizobium caulinodans nitrogen-fixation regulatory gene, nifA, is controlled by the cellular nitrogen and oxygen status. *Molecular microbiology* **3**, 825-838 (1989).
535. Poudel, S. et al. Electron transfer to nitrogenase in different genomic and metabolic backgrounds. *Journal of bacteriology* **200**, e00757-00717 (2018).
536. Boyd, E. & Peters, J.W. New insights into the evolutionary history of biological nitrogen fixation. *Front. Microbiol.* **4**, 201 (2013).
537. Shah, V.K., Stacey, G. & Brill, W.J. Electron transport to nitrogenase. Purification and characterization of pyruvate: flavodoxin oxidoreductase. The nifJ gene product. *J. Biol. Chem.* **258**, 12064-12068 (1983).

538. Schmehl, M. et al. Identification of a new class of nitrogen fixation genes in Rhodobacter capsalatus: a putative membrane complex involved in electron transport to nitrogenase. *Mol Gen Genet* **241**, 602-615 (1993).
539. Edgren, T. & Nordlund, S. The fixABCX genes in Rhodospirillum rubrum encode a putative membrane complex participating in electron transfer to nitrogenase. *Journal of bacteriology* **186**, 2052-2060 (2004).
540. Pascuan, C., Fox, A.R., Soto, G. & Ayub, N.D. Exploring the ancestral mechanisms of regulation of horizontally acquired nitrogenases. *Journal of molecular evolution* **81**, 84-89 (2015).
541. Yan, Y. et al. Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated Pseudomonas stutzeri A1501. *PNAS* (2008).
542. Kechris, K.J., Lin, J.C., Bickel, P.J. & Glazer, A.N. Quantitative exploration of the occurrence of lateral gene transfer by using nitrogen fixation genes as a case study. *PNAS* **103**, 9584-9589 (2006).
543. Thöny, B., Anthamatten, D. & Hennecke, H. Dual control of the Bradyrhizobium japonicum symbiotic nitrogen fixation regulatory operon fixR nifA: analysis of cis-and trans-acting elements. *Journal of bacteriology* **171**, 4162-4169 (1989).
544. Li, X.-X., Liu, Q., Liu, X.-M., Shi, H.-W. & Chen, S.-F. Using synthetic biology to increase nitrogenase activity. *Microb Cell Fact* **15**, 43 (2016).
545. Han, Y. et al. Interspecies Transfer and Regulation of Pseudomonas stutzeri A1501 Nitrogen Fixation Island in Escherichia coli. *J. Microbiol. Biotechnol.* **25**, 1339-1348 (2015).
546. Dixon, R.A. & Postgate, J.R. Genetic transfer of nitrogen fixation from Klebsiella pneumoniae to Escherichia coli. *Nature* **237**, 102 (1972).
547. Cannon, F., Dixon, R., Postgate, J. & Primrose, S. Chromosomal integration of Klebsiella nitrogen fixation genes in Escherichia coli. *Microbiology* **80**, 227-239 (1974).
548. Cannon, F., Dixon, R., Postgate, J. & Primrose, S. Plasmids formed in nitrogen-fixing Escherichia coli-Klebsiella pneumoniae hybrids. *Microbiology* **80**, 241-251 (1974).
549. Wang, L. et al. A minimal nitrogen fixation gene cluster from Paenibacillus sp. WLY78 enables expression of active nitrogenase in Escherichia coli. *PLOS Genetics* **9**, e1003865 (2013).
550. Temme, K., Zhao, D. & Voigt, C.A. Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca. *Proceedings of the National Academy of Sciences* **109**, 7085-7090 (2012).
551. Smanski, M.J. et al. Functional optimization of gene clusters by combinatorial design and assembly. *Nature biotechnology* **32**, 1241 (2014).
552. Zhang, L., Liu, X., Li, X. & Chen, S. Expression of the N2 fixation gene operon of Paenibacillus sp. WLY78 under the control of the T7 promoter in Escherichia coli BL21. *Biotechnol Lett* **37**, 1999-2004 (2015).
553. POSTGATE, J.R. & KENT, H.M. Qualitative Evidence for Expression of Klebsiella pneumoniae nif in Pseudomonas putida. *Microbiology* **133**, 2563-2566 (1987).
554. Setten, L. et al. Engineering Pseudomonas protegens Pf-5 for nitrogen fixation and its application to improve plant growth under nitrogen-deficient conditions. *PLoS One* **8**, e63666 (2013).
555. Dixon, R. & Kahn, D. Genetic regulation of biological nitrogen fixation. *Nature Reviews Microbiology* **2**, 621 (2004).
556. Fischer, H.-M. Environmental regulation of rhizobial symbiotic nitrogen fixation genes. *Trends in Microbiology* **4**, 317-320 (1996).
557. Tsoy, O.V., Ravcheev, D.A., Čuklina, J. & Gelfand, M.S. Nitrogen fixation and molecular oxygen: comparative genomic reconstruction of transcription regulation in Alphaproteobacteria. *Front. Microbiol.* **7**, 1343 (2016).
558. Hill, S., Kennedy, C., Kavanagh, E., Goldberg, R.B. & Hanau, R. Nitrogen fixation gene (nifL) involved in oxygen regulation of nitrogenase synthesis in K. peumoniae. *Nature* **290**, 424 (1981).
559. Mandon, K. et al. Poly-beta-hydroxybutyrate turnover in Azorhizobium caulinodans is required for growth and affects nifA expression. *Journal of bacteriology* **180**, 5070-5076 (1998).

560. Little, R., Reyes-Ramirez, F., Zhang, Y., van Heeswijk, W.C. & Dixon, R. Signal transduction to the Azotobacter vinelandii NIFL–NIFA regulatory system is influenced directly by interaction with 2-oxoglutarate and the PII regulatory protein. *The EMBO journal* **19**, 6041-6050 (2000).
561. Poole, P., Ramachandran, V. & Terpolilli, J. Rhizobia: from saprophytes to endosymbionts. *Nature Reviews Microbiology* (2018).
562. Kong, Q., Wu, Q., Ma, Z. & Shen, S. Oxygen sensitivity of the nifLA promoter of Klebsiella pneumoniae. *Journal of bacteriology* **166**, 353-356 (1986).
563. Brooks, S.J., Collins, J.J. & Brill, W.J. Repression of nitrogen fixation in Klebsiella pneumoniae at high temperature. *Journal of bacteriology* **157**, 460-464 (1984).
564. Peoples, M., Ladha, J. & Herridge, D. in Management of Biological Nitrogen Fixation for the Development of More Productive and Sustainable Agricultural Systems 83-101 (Springer, 1995).
565. Barakat, M., Cheviron, B. & Angulo-Jaramillo, R.J.A.W.M. Influence of the irrigation technique and strategies on the nitrogen cycle and budget: A review. **178**, 225-238 (2016).
566. Burger, M., Jackson, L.E.J.S.B. & Biochemistry Microbial immobilization of ammonium and nitrate in relation to ammonification and nitrification rates in organic and conventional cropping systems. **35**, 29-36 (2003).
567. MAcNEIL, D. & Brill, W.J.J.J.o.b. Mutations in nif genes that cause Klebsiella pneumoniae to be derepressed for nitrogenase synthesis in the presence of ammonium. **144**, 744-751 (1980).
568. Bali, A., Blanco, G., Hill, S., Kennedy, C.J.A. & microbiology, e. Excretion of ammonium by a nifL mutant of Azotobacter vinelandii fixing nitrogen. **58**, 1711-1718 (1992).
569. Arsène, F., Kaminski, P.A. & Elmerich, C.J.F.m.l. Control of Azospirillum brasilense NifA activity by PII: effect of replacing Tyr residues of the NifA N-terminal domain on NifA activity. **179**, 339-343 (1999).
570. Souza, E., Pedrosa, F., Drummond, M., Rigo, L. & Yates, M.J.J.o.b. Control of Herbaspirillum seropedicaeNifA Activity by Ammonium Ions and Oxygen. **181**, 681-684 (1999).
571. Paschen, A., Drepper, T., Masepohl, B. & Klipp, W. Rhodobacter capsulatus nifA mutants mediating nif gene expression in the presence of ammonium. *FEMS Microbiol Lett* **200**, 207-213 (2001).
572. Rey, F.E., Heiniger, E.K. & Harwood, C.S. Redirection of metabolism for biological hydrogen production. *Appl Environ Microbiol* **73**, 1665-1671 (2007).
573. Smanski, M.J. et al. Synthetic biology to access and expand nature's chemical diversity. *Nature Reviews Microbiology* **14**, 135 (2016).
574. Scupham, A.J. et al. Inoculation with Sinorhizobium meliloti RMBPC-2 increases alfalfa yield compared with inoculation with a nonengineered wild-type strain. *Applied and environmental microbiology* **62**, 4260-4262 (1996).
575. Fox, A.R. et al. Major cereal crops benefit from biological nitrogen fixation when inoculated with the nitrogen-fixing bacterium Pseudomonas protegens Pf-5 X940. *Environ. Microbiol.* **18**, 3522-3534 (2016).
576. Suthar, H., Hingurao, K., Vaghashiya, J. & Parmar, J. in Microorganisms for Green Revolution 229-252 (Springer, 2017).
577. Lucy, M., Reed, E. & Glick, B.R. Applications of free living plant growth-promoting rhizobacteria. *Antonie van leeuwenhoek* **86**, 1-25 (2004).
578. Mahmood, A., Turgay, O.C., Farooq, M. & Hayat, R. Seed biopriming with plant growth promoting rhizobacteria: a review. *FEMS Microbiol Ecol* **92** (2016).
579. Velusamy, P., Immanuel, J.E., Gnanamanickam, S.S. & Thomashow, L.J.C.j.o.m. Biological control of rice bacterial blight by plant-associated bacteria producing 2, 4-diacetylphloroglucinol. **52**, 56-65 (2006).
580. Mishra, J. (Department of Environmental Microbiology, School for Environmental Sciences, Babasaheb Bhimrao Ambedkar University, 2016).
581. Weller, D.M. et al. Role of 2,4-diacetylphloroglucinol-producing fluorescent Pseudomonas spp. in the defense of plant roots. *Plant Biol (Stuttg)* **9**, 4-20 (2007).

582. Kaiser, C. et al. Exploring the transfer of recent plant photosynthates to soil microbes: mycorrhizal pathway vs direct root exudation. *New Phytol* **205**, 1537-1551 (2015).
583. Mus, F. et al. Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. *Appl Environ Microbiol* **82**, 3698-3710 (2016).
584. Perrine-Walker, F.M., Prayitno, J., Rolfe, B.G., Weinman, J.J. & Hocart, C.H. Infection process and the interaction of rice roots with rhizobia. *J Exp Bot* **58**, 3343-3350 (2007).
585. Gough, C. et al. Specific flavonoids promote intercellular root colonization of Arabidopsis thaliana by Azorhizobium caulinodans ORS571. *Molecular plant-microbe interactions* **10**, 560-570 (1997).
586. O'Callaghan, K.J. et al. Effects of Glucosinolates and Flavonoids on Colonization of the Roots of Brassica napus byAzorhizobium caulinodans ORS571. *Appl Environ Microbiol* **66**, 2185-2191 (2000).
587. Mongiardini, E.J. et al. The rhizobial adhesion protein RapA1 is involved in adsorption of rhizobia to plant roots but not in nodulation. *FEMS Microbiol Ecol* **65**, 279-288 (2008).
588. Matilla, M.A., Espinosa-Urgel, M., Rodríguez-Herva, J.J., Ramos, J.L. & Ramos-González, M.I. Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. *Genome Biol* **8**, R179 (2007).
589. Casas, M.I., Duarte, S., Doseff, A.I. & Grotewold, E.J.F.i.p.s. Flavone-rich maize: an opportunity to improve the nutritional value of an important commodity crop. **5**, 440 (2014).
590. Ogo, Y., Ozawa, K., Ishimaru, T., Murayama, T. & Takaiwa, F.J.P.b.j. Transgenic rice seed synthesizing diverse flavonoids at high levels: a new platform for flavonoid production with associated health benefits. **11**, 734-746 (2013).
591. Bacilio-Jiménez, M. et al. Chemical characterization of root exudates from rice (Oryza sativa) and their effects on the chemotactic response of endophytic bacteria. **249**, 271-277 (2003).
592. Pini, F. et al. Lux bacterial biosensors for in vivo spatiotemporal mapping of root secretion. *Plant Physiol.*, pp. 01302.02016 (2017).
593. Wang, Y., Zhang, J., Sun, Y., Feng, J. & Zhang, X.J.M. Evaluating the Potential Value of Natural Product Cuminic Acid against Plant Pathogenic Fungi in Cucumber. **22**, 1914 (2017).
594. Zhou, X. & Wu, F.J.S.r. Vanillic acid changed cucumber (Cucumis sativus L.) seedling rhizosphere total bacterial, Pseudomonas and Bacillus spp. communities. **8**, 4929 (2018).
595. Soto, M.J., Sanjuan, J. & Olivares, J. Rhizobia and plant-pathogenic bacteria: common infection weapons. *Microbiology* **152**, 3167-3174 (2006).
596. Poonguzhali, S., Madhaiyan, M. & Sa, T.J.R.i.m. Quorum-sensing signals produced by plant-growth promoting Burkholderia strains under in vitro and in planta conditions. **158**, 287-294 (2007).
597. Sanchez-Contreras, M., Bauer, W.D., Gao, M., Robinson, J.B. & Downie, J.A.J.P.T.o.t.R.S.o.L.B.B.S. Quorum-sensing regulation in rhizobia and its role in symbiotic interactions with legumes. **362**, 1149-1163 (2007).
598. Chagas, F.O., de Cassia Pessotti, R., Caraballo-Rodríguez, A.M. & Pupo, M.T.J.C.S.R. Chemical signaling involved in plant–microbe interactions. **47**, 1652-1704 (2018).
599. Schikora, A., Schenk, S.T. & Hartmann, A.J.P.m.b. Beneficial effects of bacteria-plant communication based on quorum sensing molecules of the N-acyl homoserine lactone group. **90**, 605-612 (2016).
600. Westervelt, P., Bloom, M.L., Mabbott, G.A. & Fekete, F.A.J.F.M.l. The isolation and identification of 3, 4-dihydroxybenzoic acid formed by nitrogen-fixing Azomonas macrocytogenes. **30**, 331-335 (1985).
601. Collinson, S.K., Doran, J.L. & Page, W.J.J.C.j.o.m. Production of 3, 4-dihydroxybenzoic acid by Azomonas macrocytogenes and Azotobacter paspali. **33**, 169-175 (1987).
602. Ryan, P.R., Dessaux, Y., Thomashow, L.S., Weller, D.M.J.P. & Soil Rhizosphere engineering and management for sustainable agriculture. **321**, 363-383 (2009).
603. Dessaux, Y., Grandclément, C. & Faure, D.J.T.i.p.s. Engineering the rhizosphere. **21**, 266-278 (2016).

604. Bressan, M. et al. Exogenous glucosinolate produced by Arabidopsis thaliana has an impact on microbes in the rhizosphere and plant roots. *The ISME journal* **3**, 1243 (2009).
605. Badri, D.V. et al. An ABC transporter mutation alters root exudation of phytochemicals that provoke an overhaul of natural soil microbiota. *Plant Physiol.* **151**, 2006-2017 (2009).
606. Abdel-Ghany, S.E., Day, I., Heuberger, A.L., Broeckling, C.D. & Reddy, A.S. Production of phloroglucinol, a platform chemical, in Arabidopsis using a bacterial gene. *Sci Rep-Uk* **6**, 38483 (2016).
607. Mondy, S. et al. An increasing opine carbon bias in artificial exudation systems and genetically modified plant rhizospheres leads to an increasing reshaping of bacterial populations. *Molecular ecology* **23**, 4846-4861 (2014).
608. Hu, L., Feng, J., Zhang, X. & Zhang, Y.-l.J.C.J.P.S. Isolation and structure detection of fungicidal components from Cuminum cyminum seed. **9**, 330-334 (2007).
609. Wang, Y., Sun, Y., Zhang, Y., Zhang, X. & Feng, J.J.M. Antifungal activity and biochemical response of cuminic acid against Phytophthora capsici Leonian. **21**, 756 (2016).
610. Bais, H.P., Weir, T.L., Perry, L.G., Gilroy, S. & Vivanco, J.M.J.A.R.P.B. The role of root exudates in rhizosphere interactions with plants and other organisms. **57**, 233-266 (2006).
611. Rasmann, S. & Turlings, T.C.J.C.o.i.p.b. Root signals that mediate mutualistic interactions in the rhizosphere. **32**, 62-68 (2016).
612. Azaizeh, H., Marschner, H., Römheld, V. & Wittenmayer, L.J.M. Effects of a vesicular-arbuscular mycorrhizal fungus and other soil microorganisms on growth, mineral nutrient acquisition and root exudation of soil-grown maize plants. **5**, 321-327 (1995).
613. Jousset, A. et al. Plants respond to pathogen infection by enhancing the antifungal gene expression of root-associated bacteria. **24**, 352-358 (2011).
614. Seal, A.N., Pratley, J.E., Haig, T. & An, M.J.J.o.c.e. Identification and quantitation of compounds in a series of allelopathic and non-allelopathic rice root exudates. **30**, 1647-1662 (2004).
615. Carvalhais, L.C. et al. Root exudation of sugars, amino acids, and organic acids by maize as affected by nitrogen, phosphorus, potassium, and iron deficiency. **174**, 3-11 (2011).
616. Szalai, G., Janda, T.J.J.o.a. & science, c. Effect of salt stress on the salicylic acid synthesis in young maize (Zea mays L.) plants. **195**, 165-171 (2009).
617. Kong, C.-H. et al. Plant neighbor detection and allelochemical response are driven by root-secreted signaling chemicals. **9**, 3867 (2018).
618. Hayat, Q., Hayat, S., Irfan, M., Ahmad, A.J.E. & botany, e. Effect of exogenous salicylic acid under changing environment: a review. **68**, 14-25 (2010).
619. Li, B. et al. Root exudates drive interspecific facilitation by enhancing nodulation and N2 fixation. *Proc Natl Acad Sci U S A* **113**, 6496-6501 (2016).
620. Pérez-Montaño, F. et al. Rice and bean AHL-mimic quorum-sensing signals specifically interfere with the capacity to form biofilms by plant-associated bacteria. **164**, 749-760 (2013).
621. Picard, C. & Bosco, M. Genetic diversity of phlD gene from 2,4-diacetylphloroglucinol-producing Pseudomonas spp. strains from the maize rhizosphere. *FEMS Microbiol Lett* **219**, 167-172 (2003).
622. Götz-Rösch, C. et al. Influence of bacterial N-acyl-homoserine lactones on growth parameters, pigments, antioxidative capacities and the xenobiotic phase II detoxification enzymes in barley and yam bean. **6**, 205 (2015).
623. Hernández-Reyes, C., Schenk, S.T., Neumann, C., Kogel, K.H. & Schikora, A.J.M.b. N-acyl-homoserine lactones-producing bacteria protect plants against plant and human pathogens. **7**, 580-588 (2014).
624. Gao, M., Teplitski, M., Robinson, J.B. & Bauer, W.D. Production of substances by Medicago truncatula that affect bacterial quorum sensing. *Mol Plant Microbe Interact* **16**, 827-834 (2003).
625. Corral-Lugo, A., Daddaoua, A., Ortega, A., Espinosa-Urgel, M. & Krell, T. Rosmarinic acid is a homoserine lactone mimic produced by plants that activates a bacterial quorum-sensing regulator. *Sci Signal* **9**, ra1 (2016).

626. Mathesius, U. et al. Extensive and specific responses of a eukaryote to bacterial quorum-sensing signals. **100**, 1444-1449 (2003).
627. Bhattacharya, A., Sood, P. & Citovsky, V.J.M.p.p. The roles of plant phenolics in defence and communication during Agrobacterium and Rhizobium infection. **11**, 705-719 (2010).
628. Meyer, A.J., Segall-Shapiro, T.H., Glassey, E., Zhang, J. & Voigt, C.A.J.N.c.b. Escherichia coli "Marionette" strains with 12 highly optimized small-molecule sensors. 1 (2018).
629. Woodruff, L.B. et al. Registry in a tube: multiplexed pools of retrievable parts for genetic design space exploration. **45**, 1553-1565 (2016).
630. Price, M.N., Huang, K.H., Arkin, A.P. & Alm, E.J.J.G.r. Operon formation is driven by co-regulation and not by horizontal gene transfer. **15**, 809-819 (2005).
631. Price, M.N., Arkin, A.P. & Alm, E.J.J.P.g. The life-cycle of operons. **2**, e96 (2006).
632. Tian, T. & Salis, H.M.J.N.a.r. A predictive biophysical model of translational coupling to coordinate and control protein expression in bacterial operons. **43**, 7137-7151 (2015).
633. Poole, P.S., Schofiel, N.A., Reid, C.J., Drew, E.M. & Walshaw, D.L.J.M. Identification of chromosomal genes located downstream of dctD that affect the requirement for calcium and the lipopolysaccharide layer of Rhizobium leguminosarum. **140**, 2797-2809 (1994).
634. Wilson, P.W. & Knight, S.G. Experiments in bacterial physiology. (Not Found, 1947).
635. Ferri, L., Gori, A., Biondi, E.G., Mengoni, A. & Bazzicalupo, M.J.P. Plasmid electroporation of Sinorhizobium strains: The role of the restriction gene hsdR in type strain Rm1021. **63**, 128-135 (2010).
636. Selbitschka, W. et al. Characterization of recA genes and recA mutants of Rhizobium meliloti and Rhizobium leguminosarum biovar viciae. **229**, 86-95 (1991).
637. Shanks, R.M. et al. Saccharomyces cerevisiae-based molecular tool kit for manipulation of genes from gram-negative bacteria. **72**, 5027-5036 (2006).
638. Prell, J., Boesten, B., Poole, P. & Priefer, U.B.J.M. The Rhizobium leguminosarum bv. viciae VF39 γ-aminobutyrate (GABA) aminotransferase gene (gabT) is induced by GABA and highly expressed in bacteroids. **148**, 615-623 (2002).
639. Silva-Rocha, R. et al. The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. **41**, D666-D675 (2012).
640. Anderson, J. et al. BglBricks: A flexible standard for biological part assembly. **4**, 1 (2010).
641. Farasat, I. et al. Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria. **10**, 731 (2014).
642. Choi, K.-H. & Schweizer, H.P.J.N.p. mini-Tn7 insertion in bacteria with single attTn7 sites: example Pseudomonas aeruginosa. **1**, 153 (2006).
643. Jacob, G., Schaefer, J., Garbow, J. & Stejskal, E.J.J.o.B.C. Solid-state NMR studies of Klebsiella pneumoniae grown under nitrogen-fixing conditions. **262**, 254-259 (1987).
644. Peters, N. & Verma, D. Phenolic compounds as regulators of gene expression in plant-microbe interactions. *Mol. Plant-Microbe Interact* **3**, 8 (1990).
645. Mathesius, U. & Watt, M. in Progress in botany 72 125-161 (Springer, 2010).
646. Chagas, F.O., de Cassia Pessotti, R., Caraballo-Rodriguez, A.M. & Pupo, M.T. Chemical signaling involved in plant–microbe interactions. *Chemical Society Reviews* **47**, 1652-1704 (2018).
647. Clear, M.R. & Hom, E.F. The evolution of symbiotic plant-microbe signaling. *Annual plant reviews online* **2** (2019).
648. van Dam, N.M. & Bouwmeester, H.J. Metabolomics in the rhizosphere: tapping into belowground chemical communication. *Trends in plant science* **21**, 256-265 (2016).
649. Strehmel, N., Böttcher, C., Schmidt, S. & Scheel, D. Profiling of secondary metabolites in root exudates of Arabidopsis thaliana. *Phytochemistry* **108**, 35-46 (2014).
650. Badri, D.V., Weir, T.L., van der Lelie, D. & Vivanco, J.M. Rhizosphere chemical dialogues: plant–microbe interactions. *Current opinion in biotechnology* **20**, 642-650 (2009).
651. Sharifi, R., Jeon, J.-S. & Ryu, C.-M. Belowground plant–microbe communications via volatile compounds. *J Exp Bot* **73**, 463-486 (2022).

652. Hassan, S. & Mathesius, U. The role of flavonoids in root–rhizosphere signalling: opportunities and challenges for improving plant–microbe interactions. *J Exp Bot* **63**, 3429-3444 (2012).
653. González, O. et al. Non-ribosomal peptide synthases from Pseudomonas aeruginosa play a role in cyclodipeptide biosynthesis, quorum-sensing regulation, and root development in a plant host. *Microb Ecol* **73**, 616-629 (2017).
654. Ferluga, S., Steindler, L. & Venturi, V. in Secondary metabolites in soil ecology 69-90 (Springer, 2008).
655. Kumar, V., Baweja, M., Singh, P.K. & Shukla, P. Recent developments in systems biology and metabolic engineering of plant–microbe interactions. *Front Plant Sci* **7**, 1421 (2016).
656. Haldar, S. & Sengupta, S. Plant-microbe cross-talk in the rhizosphere: insight and biotechnological potential. *The open microbiology journal* **9**, 1 (2015).
657. Sanyal, D., Subhash, G.V., Dwivedi, V. & Dasgupta, S. in Climate Change and the Microbiome 695-715 (Springer, 2021).
658. Imam, J., Singh, P.K. & Shukla, P. Plant microbe interactions in post genomic era: perspectives and applications. *Front. Microbiol.* **7**, 1488 (2016).
659. Farrar, K., Bryant, D. & Cope-Selby, N. Understanding and engineering beneficial plant–microbe interactions: plant growth promotion in energy crops. *Plant biotechnology journal* **12**, 1193-1206 (2014).
660. Stephens, K. & Bentley, W.E. Synthetic biology for manipulating quorum sensing in microbial consortia. *Trends in Microbiology* **28**, 633-643 (2020).
661. Boo, A., Amaro, R.L. & Stan, G.-B. Quorum sensing in synthetic biology: a review. *Current Opinion in Systems Biology* **28**, 100378 (2021).
662. Canton, B., Labno, A. & Endy, D. Refinement and standardization of synthetic biological parts and devices. *Nature biotechnology* **26**, 787-793 (2008).
663. Voigt, C.A. Genetic parts to program bacteria. *Current opinion in biotechnology* **17**, 548-557 (2006).
664. Tsao, C.-Y., Quan, D.N. & Bentley, W.E. Development of the quorum sensing biotechnological toolbox. *Current Opinion in Chemical Engineering* **1**, 396-402 (2012).
665. Vaiana, C.A. et al. Characterizing chemical signaling between engineered "microbial sentinels" in porous microplates. *Molecular Systems Biology* **18**, e10785 (2022).
666. Collins, C.H., Leadbetter, J.R. & Arnold, F.H. Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. *Nature biotechnology* **24**, 708-712 (2006).
667. Du, P.a.Z.H.a.Z.H.a.W.R.a.H.J.a.T.Y.a.L.X.a.L.X.a.W.M.a.X.Y.a.Q. De novo design of an intercellular signaling toolbox for multi-channel cell–cell communication and biological computation. *Nature Communications* **11**, 4226 (2020).
668. Scott, S.R. & Hasty, J. Quorum Sensing Communication Modules for Microbial Consortia. *ACS Synth Biol* **5**, 969-977 (2016).
669. Chen, T., Al-Radhawi, M.A., Voigt, C.A. & Sontag, E.D. A synthetic distributed genetic multi-bit counter. *IScience* **24**, 103526 (2021).
670. Choudhary, S. & Schmidt-Dannert, C. Applications of quorum sensing in biotechnology. *Appl Microbiol Biotechnol* **86**, 1267-1279 (2010).
671. Dinh, C.V., Chen, X. & Prather, K.L. Development of a quorum-sensing based circuit for control of coculture population composition in a naringenin production system. *ACS Synth Biol* **9**, 590-597 (2020).
672. Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H. & Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130-1134 (2005).
673. Tamsir, A., Tabor, J.J. & Voigt, C.A. Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. *Nature* **469**, 212-215 (2011).
674. Tabor, J.J. et al. A synthetic genetic edge detection program. *Cell* **137**, 1272-1281 (2009).

675. Danino, T., Mondragón-Palomino, O., Tsimring, L. & Hasty, J. A synchronized quorum of genetic clocks. *Nature* **463**, 326-330 (2010).
676. Gupta, A., Reizman, I.M.B., Reisch, C.R. & Prather, K.L. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nature biotechnology* **35**, 273-279 (2017).
677. Brenner, K., Karig, D.K., Weiss, R. & Arnold, F.H. Engineered bidirectional communication mediates a consensus in a microbial biofilm consortium. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 17300-17304 (2007).
678. Kobayashi, H. et al. Programmable cells: interfacing natural and engineered gene networks. *PNAS* **101**, 8414-8419 (2004).
679. Billerbeck, S. et al. A scalable peptide-GPCR language for engineering multicellular communication. *Nature Communications* **9**, 5057 (2018).
680. Martin, S.G. Quorum sensing with pheromones. *Nature microbiology* **4**, 1430-1431 (2019).
681. Khakhar, A., Bolten, N.J., Nemhauser, J. & Klavins, E. Cell–cell communication in yeast using auxin biosynthesis and auxin responsive CRISPR transcription factors. *ACS Synth Biol* **5**, 279-286 (2016).
682. Haskett, T.L. et al. Engineered plant control of associative nitrogen fixation. *PNAS* **119**, e2117465119 (2022).
683. Fray, R.G. et al. Plants genetically modified to produce N-acylhomoserine lactones communicate with bacteria. *Nature biotechnology* **17**, 1017-1020 (1999).
684. Geddes, B.A. et al. Engineering transkingdom signalling in plants to control gene expression in rhizosphere bacteria. *Nature communications* **10**, 1-11 (2019).
685. Zhang, Z. et al. Inducible Gene Expression Systems in Plants. *Plant Gene and Trait* **11** (2020).
686. Aoyama, T. & Chua, N.H. A glucocorticoid-mediated transcriptional induction system in transgenic plants. *The Plant Journal* **11**, 605-612 (1997).
687. Beerli, R.R., Schopfer, U., Dreier, B. & Barbas, C.F. Chemically regulated zinc finger transcription factors. *J. Biol. Chem.* **275**, 32617-32627 (2000).
688. Roslan, H.A. et al. Characterization of the ethanol-inducible alc gene-expression system in Arabidopsis thaliana. *The Plant Journal* **28**, 225-235 (2001).
689. Zuo, J. & Chua, N.-H. Chemical-inducible systems for regulated expression of plant genes. *Current opinion in biotechnology* **11**, 146-151 (2000).
690. Corrado, G. & Karali, M. Inducible gene expression systems and plant biotechnology. *Biotechnology Advances* **27**, 733-743 (2009).
691. Morey, K.J. et al. Crosstalk between endogenous and synthetic components – synthetic signaling meets endogenous components. *Biotechnology Journal* **7**, 846-855 (2012).
692. Verhounig, A., Karcher, D. & Bock, R. Inducible gene expression from the plastid genome by a synthetic riboswitch. *Proc Natl Acad Sci U S A* **107**, 6204-6209 (2010).
693. Calatrava-Morales, N., McIntosh, M. & Soto, M.J. Regulation mediated by N-acyl homoserine lactone quorum sensing signals in the rhizobium-legume symbiosis. *Genes* **9**, 263 (2018).
694. Cha, C., Gao, P., Chen, Y.-C., Shaw, P.D. & Farrand, S.K. Production of acyl-homoserine lactone quorum-sensing signals by gram-negative plant-associated bacteria. *Molecular Plant-Microbe Interactions* **11**, 1119-1129 (1998).
695. Viswanath, G., Sekar, J. & Ramalingam, P.V. Detection of diverse N-Acyl homoserine lactone signalling molecules among bacteria associated with rice rhizosphere. *Curr Microbiol* **77**, 3480-3491 (2020).
696. Chong, T.-M. et al. Characterization of quorum sensing and quorum quenching soil bacteria isolated from Malaysian tropical montane forest. *Sensors-Basel* **12**, 4846-4859 (2012).
697. Rosier, A., Bishnoi, U., Lakshmanan, V., Sherrier, D.J. & Bais, H.P. A perspective on inter-kingdom signaling in plant-beneficial microbe interactions. *Plant Mol Biol* **90**, 537-548 (2016).

698. Begum, J.F., Tamilarasi, M., Pushpakanth, P. & Balachandar, D. A simple method for direct isolation of N-acyl-L-homoserine lactone mediated biofilm-forming rhizobacteria from roots. *Journal of Microbiological Methods* **156**, 34-39 (2019).
699. Steidle, A. et al. Visualization of N-acylhomoserine lactone-mediated cell-cell communication between bacteria colonizing the tomato rhizosphere. *Appl Environ Microbiol* **67**, 5761-5770 (2001).
700. Delalande, L. et al. N-hexanoyl-L-homoserine lactone, a mediator of bacterial quorum-sensing regulation, exhibits plant-dependent stability and may be inactivated by germinating Lotus corniculatus seedlings. *FEMS Microbiol Ecol* **52**, 13-20 (2005).
701. Gantner, S. et al. In situ quantitation of the spatial scale of calling distances and population density-independent N-acylhomoserine lactone-mediated communication by rhizobacteria colonized on plant roots. *FEMS Microbiol Ecol* **56**, 188-194 (2006).
702. Dulla, G. & Lindow, S.E. Quorum size of Pseudomonas syringae is small and dictated by water availability on the leaf surface. *PNAS* **105**, 3082-3087 (2008).
703. von Rad, U. et al. Response of Arabidopsis thaliana to N-hexanoyl-DL-homoserine-lactone, a bacterial quorum sensing molecule produced in the rhizosphere. *Planta* **229**, 73-85 (2008).
704. Babenko, L.M., Kosakivska, I.V. & Romanenko, K.O. Molecular mechanisms of N-acyl homoserine lactone signals perception by plants. *Cell Biology International* **46**, 523-534 (2022).
705. Mathesius, U. et al. Extensive and specific responses of a eukaryote to bacterial quorum-sensing signals. *Proc Natl Acad Sci U S A* **100**, 1444-1449 (2003).
706. Sieper, T. et al. N-acyl-homoserine lactone uptake and systemic transport in barley rest upon active parts of the plant. *New Phytol* **201**, 545-555 (2014).
707. Song, S., Jia, Z., Xu, J., Zhang, Z. & Bian, Z. N-butyryl-homoserine lactone, a bacterial quorum-sensing signaling molecule, induces intracellular calcium elevation in Arabidopsis root cells. *Biochem Biophys Res Commun* **414**, 355-360 (2011).
708. Teplitski, M., Mathesius, U. & Rumbaugh, K.P. Perception and degradation of N-acyl homoserine lactone quorum sensing signals by mammalian and plant cells. *Chem Rev* **111**, 100-116 (2011).
709. Uroz, S. & Heinonsalo, J. Degradation of N-acyl homoserine lactone quorum sensing signal molecules by forest root-associated fungi. *FEMS Microbiol Ecol* **65**, 271-278 (2008).
710. Lo, K.J., Lin, S.S., Lu, C.W., Kuo, C.H. & Liu, C.T. Whole-genome sequencing and comparative analysis of two plant-associated strains of Rhodopseudomonas palustris (PS3 and YSC3). *Sci Rep* **8**, 12769 (2018).
711. Schaefer, A.L. et al. A new class of homoserine lactone quorum-sensing signals. *Nature* **454**, 595-599 (2008).
712. Palmer, A.G. & Blackwell, H.E. Deciphering a protolanguage for bacteria-host communication. *Nat Chem Biol* **4**, 452-454 (2008).
713. Hirakawa, H. et al. Activity of the Rhodopseudomonas palustris p-coumaroyl-homoserine lactone-responsive transcription factor RpaR. *J Bacteriol* **193**, 2598-2607 (2011).
714. Du, X. et al. Rhodopseudomonas palustris Quorum Sensing Molecule pC-HSL Induces Systemic Resistance to TMV Infection via Upregulation of NbSIPK/NbWIPK Expressions in Nicotiana benthamiana. *Phytopathology* **111**, 500-508 (2021).
715. Cooley, M., Chhabra, S.R. & Williams, P. N-Acylhomoserine lactone-mediated quorum sensing: a twist in the tail and a blow for host immunity. *Chem Biol* **15**, 1141-1147 (2008).
716. Momb, J., Yoon, D.W. & Fast, W. Enzymic disruption of N-aroyl-L-homoserine lactone-based quorum sensing. *Chembiochem* **11**, 1535-1537 (2010).
717. Costa-Gutierrez, S.B. et al. Plant growth promotion by Pseudomonas putida KT2440 under saline stress: role of eptA. *Appl Microbiol Biotechnol* **104**, 4577-4592 (2020).
718. Bernal, P., Allsopp, L.P., Filloux, A. & Llamas, M.A. The Pseudomonas putida T6SS is a plant warden against phytopathogens. *ISME J* **11**, 972-987 (2017).
719. Subramoni, S. & Venturi, V. PpoR is a conserved unpaired LuxR solo of Pseudomonas putida which binds N-acyl homoserine lactones. *Bmc Microbiol* **9**, 125 (2009).

720. Vaiana, C.A. et al. Characterizing chemical signaling between engineered "microbial sentinels" in porous microplates. *Mol Syst Biol* **18**, e10785 (2022).
721. Vannini, A. et al. The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. *Embo J* **21**, 4393-4401 (2002).
722. Ford, T.J. & Silver, P.A. Synthetic biology expands chemical control of microorganisms. *Curr Opin Chem Biol* **28**, 20-28 (2015).
723. Salis, H., Tamsir, A. & Voigt, C. Engineering bacterial signals and sensors. *Bacterial Sensing and Signaling* **16**, 194-225 (2009).
724. Brophy, J.A. Toward synthetic plant development. *Plant Physiol.* **188**, 738-748 (2022).
725. Haskett, T.L. et al. Engineered plant control of associative nitrogen fixation. *Proc Natl Acad Sci U S A* **119**, e2117465119 (2022).
726. ORTÍZ-CASTRO, R., MARTÍNEZ-TRUJILLO, M. & LÓPEZ-BUCIO, J. N-acyl-L-homoserine lactones: a class of bacterial quorum-sensing signals alter post-embryonic root development in Arabidopsis thaliana. *Plant, cell & environment* **31**, 1497-1509 (2008).
727. Kodama, S., Okada, K., Inui, H. & Ohkawa, H. Aryl hydrocarbon receptor (AhR)-mediated reporter gene expression systems in transgenic tobacco plants. *Planta* **227**, 37-45 (2007).
728. Hsu, T.M. et al. Employing a biochemical protecting group for a sustainable indigo dyeing strategy. *Nature chemical biology* **14**, 256-261 (2018).
729. Fahad, S. et al. Phytohormones and plant responses to salinity stress: a review. *Plant growth regulation* **75**, 391-404 (2015).
730. Kan, J., Fang, R. & Jia, Y. Interkingdom signaling in plant-microbe interactions. *Science China Life Sciences* **60**, 785-796 (2017).
731. Bauer, W.D. & Robinson, J.B. Disruption of bacterial quorum sensing by other organisms. *Current opinion in biotechnology* **13**, 234-237 (2002).
732. Ahlgren, N.A., Harwood, C.S., Schaefer, A.L., Giraud, E. & Greenberg, E.P. Aryl-homoserine lactone quorum sensing in stem-nodulating photosynthetic bradyrhizobia. *PNAS* **108**, 7183-7188 (2011).
733. Geske, G.D. et al. Comparative analyses of N-acylated homoserine lactones reveal unique structural features that dictate their ability to activate or inhibit quorum sensing. *ChemBioChem* **9**, 389-400 (2008).
734. Delépine, B., Duigou, T., Carbonell, P. & Faulon, J.-L. RetroPath2. 0: A retrosynthesis workflow for metabolic engineers. *Metabolic Engineering* **45**, 158-170 (2018).
735. Lin, G.-M., Warden-Rothman, R. & Voigt, C.A. Retrosynthetic design of metabolic pathways to chemicals not found in nature. *Current Opinion in Systems Biology* **14**, 82-107 (2019).
736. Collins, C.H., Arnold, F.H. & Leadbetter, J.R. Directed evolution of Vibrio fischeri LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones. *Molecular microbiology* **55**, 712-723 (2005).
737. Hawkins, A.C., Arnold, F.H., Stuermer, R., Hauer, B. & Leadbetter, J.R. Directed evolution of Vibrio fischeri LuxR for improved response to butanoyl-homoserine lactone. *Appl Environ Microbiol* **73**, 5775-5781 (2007).
738. Tashiro, Y. et al. Directed evolution of the autoinducer selectivity of Vibrio fischeri LuxR. *The Journal of General and Applied Microbiology* **62**, 240-247 (2016).
739. Li, S.-Z. et al. Influence of the d/l configuration of N-acyl-homoserine lactones (AHLs) and analogues on their Lux-R dependent quorum sensing activity. *Bioorganic Chemistry* **77**, 215-222 (2018).
740. Shrestha, A. & Schikora, A. AHL-priming for enhanced resistance as a tool in sustainable agriculture. *FEMS Microbiol Ecol* **96**, fiaa226 (2020).
741. Chae, D.H., Kim, D.R., Cheong, M.S., Lee, Y.B. & Kwak, Y.S. Investigating the Induced Systemic Resistance Mechanism of 2,4-Diacetylphloroglucinol (DAPG) using DAPG Hydrolase-Transgenic Arabidopsis. *Plant Pathol J* **36**, 255-266 (2020).

742. Kwak, Y.S. et al. Diversity, virulence, and 2,4-diacetylphloroglucinol sensitivity of Gaeumannomyces graminis var. tritici isolates from Washington state. *Phytopathology* **99**, 472-479 (2009).
743. Persad-Russell, R. et al. Specific Bacterial Pathogen Phytosensing Is Enabled by a Synthetic Promoter-Transcription Factor System in Potato. *Front Plant Sci* **13**, 873480 (2022).
744. Mazarei, M., Teplova, I., Hajimorad, M.R. & Stewart, C.N. Pathogen Phytosensing: Plants to Report Plant Pathogens. *Sensors (Basel)* **8**, 2628-2641 (2008).
745. Kovalchuk, I. & Kovalchuk, O. Transgenic Plants as Sensors of Environmental Pollution Genotoxicity. *Sensors (Basel)* **8**, 1539-1558 (2008).
746. Lindsey, B.E., Rivero, L., Calhoun, C.S., Grotewold, E. & Brkljacic, J. Standardized Method for High-throughput Sterilization of Arabidopsis Seeds. *J Vis Exp* (2017).
747. Bauer, W.D. & Mathesius, U. Plant responses to bacterial quorum sensing signals. *Curr Opin Plant Biol* **7**, 429-433 (2004).